רשות הפטנטים

מדינת ישראל
משרד המשפטים

This is to certify that annexed hereto is a true copy of the documents as originally deposited with the patent application of which particulars are specified on the first page of the annex.

זאת לתעודה כי רצופים בזה העתקים נכונים של המסמכים שהופקדו לכתחילה עם הבקשה לפטנט לפי הפרטים הרשומים בעמוד הראשון של הנספח.

This 27-07-2010 היום

רשם הפטנטים

Commissioner of Patents

ד"ר מאיר נועם

רשם הפטנטים, המדגמים וסימני המסחר

מאושר
Certified

State of Israel
Patent Office

מדינת ישראל
רשות הפטנטים

Patent Specification פירוט פטנט

[11] Application No: **154560** **מספר בקשה:**

[22] Date of application: **20/02/2003** **תאריך הבקשה:**

[54] COMMUNICATION SYSTEM AND METHOD FOR MEDIA ACCESS CONTROL

מערכת תקשורת ושיטה לבקשת גישה לערוץ

ארגז 452881 | תיק 154560

[71] **Patentee (s):**	**בעל (י) הפטנט:**
IAMBA TECHNOLOGIES LTD. Customer No: 769831 MILLENNIUM BUILDING 3 HATIDHAR STREET. RANANNA Israel	איימבה טכנולוגיות בע"מ קוד לקוח: 769831 בניין המילניום רח' התידהר 3 רעננה 43000

Address for service: המען למסירת הודעות:

[74] OREN RECHES
17 HATIDHAR,
BEIT HATAMAR,
P.O.B. 2213,
RAANANA

אורן רכס
התדהר 17
ת.ד. 2213
רעננה 43665

[48] Date of Publication: 20/01/2008 תאריך פרסום:

COMMUNICATION SYSTEM AND METHOD FOR MEDIA ACCESS CONTROL

FIELD OF THE INVENTION

The present invention relates to communication systems and methods in general, and to methods and systems for media access control of a shared communication media. Even more particularly, to methods and systems for controlling upstream (from a network element to a Headend that is coupled to multiple network elements) transmission of a point to multipoint optical network. The invention is applicable to, but not limited to, point to multipoint passive optical networks.

BACKGROUND OF THE INVENTION

Optical access networks, such as point to multipoint optical access networks are known in the art. Point to multipoint optical access networks, such as passive optical networks allow to couple a Headend to multiple network units, thus allowing downstream transmission from the Headend to the multiple network units (point to multipoint) and allowing upstream transmissions, whereas upstream transmission is made from a single network unit to the Headend and is also referred to as point to point transmission.

ITU-T Recommendation G.983.1 defines an access network that utilizes optical fiber technology to convey point to multipoint downstream traffic from a Headend such as an Optical Line Termination (OLT) to multiple network units such as Optical Network Units (ONUs) or Optical Network Terminations (ONTs) and to convey upstream traffic from ONUs and/or OLTs to the OLT.

An ONU provides, either directly or remotely the user-side interface of the Optical Access Network (OAN).

An OLT provides the network side interface of the OAN and is connected to at least one Optical Distribution Network (ODN).

An ODN provides the optical transmission means between the OLT and the ONUs. An ODN usually includes passive optical components such as optical fibers, optical connectors, passive branching components, passive optical attenuators and splices.

The optical access network supports Asynchronous Transfer Mode (ATM) based data transmission. ATM based data transmission allows to support more than a single class of service. When a packet, such as but not limited to Internet Protocol (IP) or Ethernet packet, is received at an ATM supporting network, it is segmented to provide a group of at least one ATM cell. Each ATM cell is routed across the ATM based network. Before exiting the ATM based network the IP packet is reassembled from the at least one ATM cell.

Each ATM cell is 53 bytes long and includes a 5-byte header and 48-byte payload. The ATM header is utilized for routing ATM cells across the ATM based network.

The optical access network can be Ethernet PON, in this case Ethernet packet are transmitted over the PON and no Segmentation and reassembly is needed, other types of communications like Ethernet, IP, GFP, TDM can be used over the PON, this invention is intended to define a way to allocate the shared resource of a network system with shared Bandwidth resources and is not limited to any specific protocol.

As multiple network units, such as ONUs or ONTs share the same media, the OLT controls upstream transmission, by assigning upstream timeslots, by implementing a Media Access Control (MAC) scheme. According to the ITU-T Recommendation G.983.4 upstream bandwidth may be assigned in three manners – (a) in response to the utilization of upstream bandwidth by each of the ONUs (or ONTs), (b) in response to upstream status reports from the ONUs or ONTs, or (c) static bandwidth allocation. More specifically, each

ONU (or ONT) can include at least one Transmission Container (T-CONT), each T-CONT has at least one queue. An ONU (or ONT) reports the queue length of all the queues in all T-CONTs that belong to him.

The OLT defines how the ONUs (or ONTs) selectively report their T-CONT's queues lengths. These reports are transmitted upstream in mini-slots assigned by the OLT or in special purpose cells (or packets). For example, An ONU (or ONT) can report the queues lengths from the most congested T-CONT, from each T-CONT in turn equally, but this is not necessarily so. In other words, the OLT defines the content of each mini slot in advance; and than only gives a grant for that mini slot.

Upstream traffic is arranged in an upstream frame of 53 timeslots (in the case of upstream data rate of 155Mbps) or 212 timeslots (for 622 Mbps). Each timeslot consists of three-bytes of PON layer overhead and either an ATM cell or a PLOAM cell.

The OLT allocates upstream bandwidth, according to the T-CONTs queue length, to bandwidth utilization or in accordance to a predefined statistical scheme and then transmits downstream data grants in downstream PLOAM cells. Assuming that the upstream and downstream bit rate are 155 Mbit/sec, then during a downstream frame of 56 cells (in 155Mbit downstream frame there are 56 timeslots/ATM cells, in upstream frame only 53), two PLOAM cells are utilized for providing 53 data grants, corresponding to the 53 timeslots within each upstream frame. When the upstream data rate is much smaller than the downstream data rate, some PLOAM cells may be empty.

In g.983.4 The T-CONTs are classified to five types, each type is characterized by an assigned bandwidth out of the following five assigned bandwidth types: (i) fixed bandwidth, (ii) assured bandwidth, (iii) non-assured bandwidth, (iv) best effort bandwidth, and (v) maximum bandwidth. It is noted that the first four assigned bandwidths are listed according to their priority, starting with the highest priority assigned bandwidth. Accordingly, the

assignment of bandwidth starts by assigning bandwidth to fixed bandwidth. The assignment limits the amount of cell transfer delay and delay variation. Assured bandwidth is assigned using the remaining bandwidth. Assured bandwidth means that a predefined average (long- range) bandwidth is assigned. It is noted that the amount of allocated bandwidth per frame can fluctuate. The yet remaining bandwidth (also referred to as surplus bandwidth) is utilized for the lower priority bandwidth assignments such as the non-assured bandwidth and the best effort bandwidth.

A type 1 T-CONT is characterized by fixed bandwidth only. Bandwidth is allocated to a type 1 T-CONT regardless of whether its queues are empty or not.

A type 2 T-CONT is characterized by assured bandwidth only.

A type 3 T-CONT is characterized by assured bandwidth and non-assured bandwidth. A type 3 T-CONT shall be allocated bandwidth equivalent to its Assured bandwidth, only when it has cells at a rate equivalent to Assured bandwidth or more than Assured bandwidth. Non-assured bandwidth shall be allocated across T-CONTs with Assured bandwidth, and by requesting surplus bandwidth in proportion to the Assured bandwidth of the individual T-CONT on the PON, e.g., Weighted Round Robin method, as surplus bandwidth. The sum of the assured bandwidth and non-assured bandwidth allocated to this T-CONT should not exceed its maximum bandwidth, which is pre-provisioned.

A type 4 T-CONT is characterized by best effort bandwidth only and does not have any guaranteed bandwidth. A type 4 T-CONT shall only use bandwidth that has not been allocated as fixed bandwidth, assured bandwidth and non-assured bandwidth to other T-CONTs sharing the same upstream bandwidth. Best-effort bandwidth is allocated to each type 4 T-CONT equivalently, e.g., based on Round Robin method, up to their predefined maximum bandwidth.

A type 5 T-CONT is the super set of type 1 – type 4 T-CONTs. Accordingly they can be characterized by at least one of the following assigned bandwidths: fixed bandwidth, assured bandwidth, non-assured bandwidth and best-effort bandwidth. It is noted that the bandwidth allocation cannot exceed the maximum bandwidth of the T-CONT.

It is noted that a T-CONT may have a priority control mechanism and/or an internal schedule that are operable to determine from which class of service queue to transmit an ATM cell in response to a data grant.

It is further noted that each assigned bandwidth type may be associated with its own class of service. Accordingly, the various mentioned above type of bandwidth assignment are associated with various class of service

U.S. patent number 5,926,478 of Ghaibeh et al. titled "Data transmission over a point-to-multipoint optical network" describes a data transmission protocol for use in an ATM-based point-to-multipoint passive optical network interconnecting a Headend facility and a plurality of network units. The Headend facility controls the upstream transmission from the network units in response to ATM cell queue sizes at the network units and in response to a selected set of service priorities. A network unit can include various queues, such as a CBR queue, a VBR queue and a ABR queue. The sizes of these queues are included within an upstream bandwidth request.

At the Ghaibeh patent downstream data is transmitted in serial data frames comprising one hundred eighty, fifty-four byte downstream slots, including two framing slots and one hundred seventy-eight ATM cell slots. Each downstream frame slot includes a one byte MAC overhead header field for transmitting upstream transmission permits allocated over twenty bit permit fields, for a total of seventy-two upstream permits allocated per downstream frame. The downstream frames are transmitted every 125 .mu.sec for an overall downstream bit rate of 622.08 Mbps. Upstream data is transmitted from an individual network units in five hundred forty bit upstream data slots, each

upstream slot having a preamble portion and a payload portion, i.e., with seventy-two upstream slots are transmitted every 125 .mu.sec, thereby forming upstream frames received at the Headend at a data rate of 311.04 Mbps.

End users negotiate with service providers to determine a class of service or service level. A service Level Agreement (SLA) defines traffic parameters from the end-user, throughout at least one network that interconnect the end-user with other end-users or with other service providers. The traffic parameters include overall delay, delay fluctuations, bandwidth allocation and the like. Many users generate and receive variable length packets, such as Internet Protocol (IP) packets or Ethernet frames. In such cases the SLA relates to the transmission of the variable length packets, and does not necessarily comply with the transmission and routing of fixed size cells originating from the packets.

The size of the variable length packet might be smaller then the aggregate size of fixed sized cells originated from the variable length packet. When the fixed sized cells are ATM cells, as a header of 5 bytes is added to each 48 bytes of IP packet, furthermore up to 47 "padding" bytes may be added for compensating for a difference between the length of the packet and the aggregate length of ATM cell payload (which is a multiplication of 48 bytes).

Assuming that the size of the packet (including AAL5 encapsulation as specified in RFC 1483 and RFC 2684 like length, CRC, etc) is $S1 = A * 48 + B$ ($0 <= B < 48$), then the aggregate size of the ATM cells that originate from the packet (S2) equals: $S2 = 53 * (\text{trunc}\{S1/ 48 \text{ byte}\} + 1) = 53*(A+1)$. It is noted that the (A+1)'th ATM cell includes B bytes originating from the IP packet and (48 - B) padding bytes. The utilization of the ATM network equals S1/S2. S1 is also referred to as the length of the relevant payload and (S2-S1) is also referred to as the overhead signals. ATM network arbitration and scheduling schemes are based upon the overall aggregate size of ATM cells (including header and stuffing bit) that are stored within queues, and not according to the aggregate "net" payload of the cells and neither upon the length of each group.

Furthermore, ATM policing schemes, that determine when to discard incoming traffic in response to various parameter such as traffic load are also cell oriented. Accordingly, most of the ATM cells originating from the same IP packet can be routed across an ATM network just to see that one ATM cell was discarded, thus all the ATM cells must be re-routed across the network.

As ATM cells are routed on a cell by cell basis, in the presence of ATM cells from many sources, consecutive cells originating from the same IP packet are not routed in a consecutive manner, thus increasing the overall delay and the delay jitter across the ATM network, and require extensive allocation of memory resources, since the actual latency of a packet is determined by the latency of the last cell of the packet.

There is a need for a system and method for improved bandwidth utilization.

There is a need for a system and method for a media access control method and controller that are based upon relevant payload.

SUMMARY OF THE PRESENT INVENTION

According to one embodiment of the invention the term "data unit information" reflects a parameter of at least one group of data cells to be transmitted over a network. The group is usually stored in at least one queue that is coupled to the network prior to the transmission. The data cells that form a single group are preferably stored in a consecutive manner within a single queue. Conveniently, each group of data cells may include relevant payload and overhead signals. Preferably, data unit information reflects the length of the relevant payload of at least one group, or the time of arrival, but this is not necessarily so. Usually, each group is associated with a single data unit information unit.

According to another embodiment of the invention the term "data unit information" reflects a parameter of a variable length packet or frame that is transmitted over the network, preferably without being segmented. The variable length packet or frame may include relevant payload and overhead signals. Preferably, data unit information reflects the length of the relevant payload of the variable packet or frame, or the time of arrival of said packet or frame, but this is not necessarily so.

The invention is based upon a first observation that data unit based media access control is more efficient than media access control schemes that are responsive to the queue length of network elements that share the same media. Data unit based media access control schemes are based upon at least one parameter of data units that are upstream and/or downstream transmitted over the network. A typical parameter may be a length of a packet that is to be transmitted over the network, or an amount of fixed sized cells that form a fixed sized cell group that has to be transmitted as a whole over the network. The group of fixed sized cells may originate from a single (variable size) packet, but this is not necessarily so. Other parameters may reflect timing information such as time of arrival of the data unit, but this is not necessarily so.

According to an aspect of the invention the network units are operable to receive variable sized packets and convert them to groups of fixed sized blocks such as but not limited to ATM cells, to be transmitted over a passive optical network.

According to another aspect of the invention the network units are operable to receive variable sized packets and to transmit them over the passive optical network. In such a case segmentation to fixed sized cells is not required, although the format of the variable sized packets may be changed before being transmitted over the passive optical network.

The invention provides a MAC scheme that is responsive to data unit information thus allowing an allocation of bandwidth according to net data on

the channel. This MAC scheme is more fair and provides compliance with Service Level Agreements more efficiently than MAC schemes that are responsive to fixed sized cells, such as the .G.983.4 MAC scheme.

The invention provides a communication system that includes an optical communication network, interconnecting a Headend (such as but not limited to a OLT) and a plurality of network units; wherein the Headend has a media access controller for issuing data grants and grants for upstream transmission of data unit information; wherein a data grant being issued at least partially in response to previously received data unit information. At least some network units out of the plurality of network units are operable to: (i) receive data to be upstream transmitted to the Headend (ii); upstream transmit data unit information associated with the received data in response to data grants issued by the media access controller (iii); and upstream transmit data to the Headend in response to data grants issued by the media access controller. According to one embodiment of the invention a data grant or a sequence of data grants authorizes an identified network unit out of the plurality of network units to upstream transmit a group of consecutive data cells during at least one consecutive timeslot. According to another aspect of the invention the system is adapted to transmit variable length packets and/or frames without performing segmentation thus the grant includes timing information defining a variable length window for upstream transmission of said packet. It is noted that according to yet a further embodiment of the invention the system is adapted to upstream transmit both fixed sized cells and variable sized packets or frames.

According to an aspect of the invention the optical communication network is a passive optical network, the Headend is an OLT and the network units are ONTs and/or ONUs. A single ONT or a single ONU can include a variety of T-CONTs. Conveniently, the communication network is ITU-T Recommendation G.983.4 compliant, and includes a segmentation and reassembly units for converting between IP packets and groups of ATM cells. Preferably, a group of

ATM cells originate from a single IP packet , and the data unit information reflects the length of the IP packet.

According to another aspect of the invention the optical communication network is a passive optical network with OLT and ONUs (or ONTs) where the data is native Ethernet, IP, GFP, ATM, or any other packet formats or combination of the above. In this case the SAR unit is not always necessary and is optional. The reports from the end units will be based on the information on the original data and the bandwidth allocation will be based on the packet information (for example length) with consideration for the transport protocol and the original data protocol.

The media access controller can include a plurality of arbitrators. Each arbitrator may be associated with a class of service out of a plurality of classes of service. At least one class of service is characterized by at least one of the following parameters: minimum latency, minimum bandwidth, maximum bandwidth, maximum burst size, minimum drop, and minimum jitter. At least one class of service may be compliant with a standard class of service, whereas the standard class of service is Diffserv class of service, IntServ class of service or MPLS class of service.

The Headend may be operable to transmit data to network units in consecutive fixed length frames; wherein each frame further includes at least one data grant. Each frame may further include a data unit information request. Usually, each frame includes a plurality of fixed length slots. The invention provides a system wherein data unit information includes data unit information units. According to another aspect of the invention the frames or packets are characterized by a non-fixed length and their transmission period is also not fixed.

It is further noted that the queue may store a packet, whereas the segmentation of a packet (necessary only according to one of the embodiments

of the invention) to at least one cell may occur only after the upstream transmission is approved.

The invention provides a system wherein the media access controller is operable to determine an amount of data unit information to be sent from a network unit. The determination may be responsive to data unit information previously transmitted from the network unit and to a data threshold, but can also be responsive to an estimation of data unit information relating to information that is yet to sent upstream from the network units. The data threshold conveniently reflects a maximal amount of data that can be upstream transmitted from the network unit to the Headend during a predefined time period. Conveniently, the predefined time period equals a MAC cycle.The invention provides a system wherein at least some of the network units are not operable to generate data unit information. In such a case the media access controller estimates data unit information relating to data to be upstream transmitted from these network units. For example, by computing the maximum possible payload for this unit data or by monitoring the received data and detecting idle cells or idle patterns.

The invention provides a system wherein at least some network units out of the plurality of network units include (i) a first input port for receiving variable length data packets and (ii) a segmenting and data unit information unit that is operable to (ii.a) segment a received variable length data packet to provide a group of fixed sized data cells, and to (ii.b) generate data unit information reflecting a parameter of the group of fixed sized cells. Conveniently, at least some network units further include a classifier, for classifying incoming data packets in response to their class of service. Typically, the variable length data packets are Internet Protocol packets or Ethernet frames and the fixed sized cells are Asynchronous Transfer Mode cells.

According to an embodiment of the invention the segmentation and reassembly unit are not necessary , such as in the case of a transmission of

variable length packets and/.or frames across the network. For example, these units may not be required at Ethemet PON compliant networks where native Ethernet packets are transferred from the ONUs (or ONTs) to the OLT, or GPON (Gigabit PON) where native Ethemet, GFP, ATM and TDM traffic is carried over the PON).

According to one aspect of the invention the invention provides a system wherein at least some network units out of the plurality of network units include (i) a second input port for receiving fixed sized cells, and (ii) an assembly unit for grouping the fixed sized cells to fixed sized cell groups. According to another aspect of the invention the network units do not segment received variable length packets or frames. Conveniently, at least some network unit further include a data unit information generator, for generating data unit information representative of a parameter of an group of fixed sized cells or ,according to another aspect of the invention representative of a parameter of a variable sized packet or frame. It is noted that the data unit information may be generated by extracting it from received data streams.

The invention provides a system wherein the media access controller is operable to provide data grants in response to at least one arbitration scheme. Conveniently, each arbitrator arbitrates between transmission requests of the same class of service, but this is not necessarily so and an arbitrator can arbitrate between transmission requests of more that a single class of service. . Issued transmission requests are queued and are selectively fetched from the queues to be downstream transmitted to the network units. Preferably, one arbitrator, such as a type 1 T-CONT arbitrator, allocates data grants in a fixed manner. Other arbitrators allocate data grants in response to data unit information, to a transmission current credit, and to a class of service rules that are defined by Service Level Agreements.

The invention provides a media access controller for controlling an access of a plurality of network units to a shared upstream channel, the media access

controller being coupled to a receiver, for receiving data unit information from the plurality of network units. Data unit information reflects at least one parameter of fixed sized cell groups to be upstream transmitted over the shared upstream channel. The media access controller includes: (i) at least one arbitration unit, coupled between the receiver and a grant allocation unit, for arbitrating between requests to upstream transmit fixed sized cell groups; and (ii) a grant allocation unit, for selecting data grants authorizing an upstream transmission of a group of fixed sized cells in response to the arbitration. According to another embodiment of the invention the data unit information reflects at least one parameter of a variable length packet or frame to be transmitted over the shared upstream channel. The at least one arbitration unit is operable to arbitrate between requests to upstream transmit variable sized packets or frames and the grant allocation unit, is able to select data grants authorizing an upstream transmission of a variable length packet or frame in response to the arbitration.

Conveniently, the grant allocation unit is operative to receive allocated data grants from the at least one arbitrating unit and to select data grants in response to a predefined priority between the at least one arbitration unit. The priority is usually dependent upon the class of service. For example, the bandwidth allocation starts by allocating bandwidth (e.g.- selecting transmission requests) of type 1 T-CONTs, and ends by allocating bandwidth to type 4 T-CONTs.

The invention provides a method for allocating upstream bandwidth of a shared upstream channel of an optical network, the optical network interconnecting a Headend with a plurality of network units, the method including the steps of:

(i) Determining data unit information to be upstream transmitted from at least one network unit.

(ii) Receiving upstream transmitted data unit information.

(iii) Issuing data grants authorizing an identified network unit to transmit upstream data in response to previously received data unit information.

Conveniently, the step of issuing includes the steps of: arbitrating between requests to transmit groups of fixed sized data cells (or transmit variable sized packets or frames, according to another embodiment of the invention); allocating data grants in response to the arbitrating; and selecting between the allocated data grants.

Conveniently, the step of arbitrating including performing at least two arbitration cycles; and wherein the step of selecting is responsive to a predefined priorities assigned to arbitration cycles.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be understood and appreciated more fully from the following detailed description taken in conjunction with the drawings in which:

Figures 1A - 1D are schematic illustration of network units, in accordance with embodiments of the invention;

Figure 2 is a schematic illustration of a Headend, in accordance with an embodiment of the invention;

Figure 3A is a schematic illustration of an IP packet and a group of three AAL5 ATM data cells being generated from the IP packet, in accordance with an embodiment of the invention;

Figure 3B is a schematic illustration of downstream data frames, in accordance with an embodiment of the invention;

Figure 4 is a schematic illustration of a portion of a queues database, in accordance with an embodiment of the invention;

Figures 5A – 5B are schematic illustrations of data unit information request vectors, in accordance with embodiments of the invention;

Figure 5C is a schematic illustration of an upstream frame, in accordance with an embodiment of the invention; and

Figures 6A – 6C are flow charts illustrating various arbitration mechanisms, in accordance with embodiments of the invention.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

The Network Unit

For convenience of explanation only, the following description mainly relates an embodiment of the invention in which fixed sized cells are transmitted. Accordingly, variable sized packets are segmented and converted to fixed sized cells, and the data unit information relates to groups of data cells. It is noted that according to another embodiment of the invention the segmentation is not required and the data unit information related to variable length packets and/.or frames that are transmitted without being segmented.

Reference is made to Figure 1A, which is a schematic illustration of a network unit NU 8, in accordance with an embodiment of the invention. A network unit can be an ONU or an ONT but this is not necessarily so.

For convenience of explanation it is assumed that fixed sized cells are ATM compliant and variable length packets are IP compliant (or Ethernet frames), but this is not necessarily so and various communication protocols and standards may be applied.

It is noted that although the arbitration schemes depend at least partially upon data unit information, and not upon queues length, queues are still the elementary unit for exchanging information In the sense that the MAC scheme determines from which queue to receive a data unit and from which queue to receive data unit information that related to data units that are stored within that

queue. Furthermore, a Headend emulates the group of cell within each queue. Accordingly, the logical and even physical location of a queue, and especially the division between T-CONTs, ONUs and ONTs is of minor significance. For convenience of explanation it is assumed that each network unit is an ONU, that each ONU has four queues, a queue for each class of service, and that classes of service that are manageable without data unit information (such as type 1 T-CONTS) are not illustrated. Furthermore, a group of queues within an ONU, an ONT or T-CONT can be viewed by the OLT as a single queue, as long as the ONU, ONT or T-CONT has a mechanism for determining from which of its queues to fetch a group of cells in response to data grants. A T-CONT in a system which is tailored to this algorithm has a single queue and this mechanism is not required It is further assumed, for convenience of explanation, that a single Headend, such as Headend 38 of Figure 2, is coupled, via a passive optical network, to thirty two NUs such as NU 8 of Figure 1A or 1B. Media access control may not be based upon data unit information regarding a specific queue when the bandwidth is allocated regardless the presence or the parameters of data (such as in the case of type 1 T-CONT queues) or in the case of NUs that are not operable to report data unit information.

NU 8 of figure 1A is operable to (a) receive variable length packets such as IP packets, via input 11 of data unit information generator 16, (b) segment the variable length packets to fixed sized cells, such as ATM cells, (c) store the ATM cells and upstream transmit the ATM cells over the passive optical network. (If is noted that the segmentation unit can be omitted according to other embodiment of the invention, for example in EPON compliant systems in which variable sized packets are transmitted over the PON and data unit information include the packets information. NU 8 is further adapted to generate and transmit data unit information. NU 8 is further operable to receive fixed sized cells via input 7 of data unit information generator 16, and perform step (c) accordingly.

NU 9 of figure 1b is operable to (a) receive fixed sized cells, such as ATM cells, via input 7 of reassembly unit 25 (b) reassemble the ATM cells to a variable length packets such as IP packets, (c) store the IP packets, (d) segment IP packets to be transmitted as ATM cells and transmit the ATM cells over the passive optical network. NU 9 is further adapted to generate and transmit data unit information. NU 9 is further operable to receive variable length packets via input 11 of data unit information generator 21, and perform steps (c) and (d) respectively.

It is noted that NU 8 and NU 9 may be adapted to transmit variable length packets over a passive optical network by omitting the segmentation of the variable length packets (either received as variable length packets as in the case of NU 8 or reassembled variable length packets as in the case of NU 9). In such a case NU 8 can store variable length packets and not fixed sized cells, and upstream assembler and transmitter 28 of Figure 1A and upstream assembler and transmitter and segmenting and intermediate queue unit of figure 1B will be altered accordingly.

It is noted that Each NU monitors the data unit information of each of the NU queues and tracks the previously upstream transmitted data unit information. Accordingly, each NU is able to determine which data unit information to prepare for upstream transmission. Data unit information generator 16 is operable to receive via input port 11 data streams that are not associated with data unit information and data streams that are associated with data unit information. The data unit information can be embedded within the data streams. In the case of transmitting cells over the passive optical network the data unit information may indicate that a group of at least one ATM cells (usually having the same Virtual Circuit Identifier and Virtual Path Identifier values) are to be routed as a group, preferably during consecutive upstream slots. A group of ATM cells may be generated by segmenting a single IP packet, but this is not necessarily so. The data unit information may reflect various parameters of the

group of ATM cells, such as aggregate "net" ATM payload size (e.g. - not including "idle" or stuffing signals), amount of ATM cells and/or a ratio between the "net" ATM payload and the aggregate size of the ATM cells of the group and/or time of arrival. It is noted that other parameters may be provided. In the case of transmission of variable length packets, such as IP packets over the passive optical network the data unit information may indicate the length of the IP packet, its time of arrival to the network unit or other information embedded within the IP packet header.

It is noted that the data unit information can be compressed by implementing various compression schemes, such as lossless compression schemes, lossy compression schemes, including quantization schemes, Haffman encoding and the like.

Data unit information generator 16 is further operable to: (a) generate ATM cells, (b) generate data unit information associated with the ATM cells, and (c) determine the class of service of the ATM cells. Data unit information generator 16 then sends the data unit information to a queues data unit information database, determines the IP packet class and provides the ATM cells to cell distributor 18 which distributes the ATM cells among data queues DQ(1,1) – DQ(1,4) 20 – 26 according to their class of service. ATM cells of the first class of service are stored at DQ(1,1) 20, ATM cells of the second class of service are stored at DQ(1,2) 22, ATM cells of the third class of service are stored at DQ(1,3) 24, and ATM cells of the fourth class of service are stored at DQ(1,4) 26. It is noted that the amount of queues and the allocation of ATM cells among the queues can be adjusted to support various class of services and various arbitration schemes, including hierarchical arbitration schemes and the like.

If, for example, the incoming data streams are ATM data streams that include ATM cells that were generated from IP packets but are not associated with data unit information then data unit information generator 16 reassembles

IP packets from the ATM cells, generates data unit information reflecting the size of the IP packet, segments the IP packets to ATM cells, provides the data unit information to data unit information database 30, determines the ATM cell class of service, and provides the ATM cells to cell distributor 18 which distributes the ATM cells among data queues DQ(1,1) – DQ(1,4) according to their class of service. Conveniently, the data unit information is also embedded within the group of ATM cells, such as in the case of AAL5 compliant ATM cells.

At this point, queues data unit information database 30 reflects the size and location of each group of ATM cells within queues DQ(1,1) – DQ(1,4), and the queues store ATM cells of up to four class of services.

Downstream receiver 34 is operable to: (i) Receive downstream traffic being transmitted over a passive optical network. The downstream traffic includes downstream broadcast data and grants. (ii) Extract the downstream data and provides it to devices/ interfaces (not shown) that are positioned in a downstream path. (iii) Extract data grants. Data grants indicate when (during which at least one consecutive selected timeslot) to transmit upstream data from a queue out of DQ(1,1) – DQ(1,4). The extracted data grants are provided to upstream transmitter 28 that in response triggers a provision of a group of ATM cells from a queue to the passive optical network, during the at least one consecutive selected timeslot. It is noted that according to one embodiment of the invention the timeslots have a fixed size (usually responsive to the transmission period of a fixed sized cell). According to another aspect of the invention variable sized packets and/or frames are transmitted and grants have to indicate a variable sized time window during which the variable sized packets and/or frame are to be transmitted. (iv) Extract data unit information grants. Data unit information grants indicate when to transmit data unit information and what data unit information to transmit. It is noted that data unit information grants may further include timing information for determining the timing of upstream

transmission (which slot and within that slot), as a plurality of NUs may transmit their data unit information during a single timeslot. The extracted data unit information grants and data grants are provided to upstream assembler and transmitter 28 that in response triggers a provision of a group of ATM cells from a queue to the passive optical network, during the at least one consecutive selected timeslot (or transmit a mini slot or any other type of messages that contain data unit information).

Reference is made to Figure 1B, which is a schematic illustration of a network unit NU 9, in accordance with an embodiment of the invention.

NU 9 includes a data unit information generator 21 that has a first input 11 for receiving variable size packets such as IP packets, packet distributor 19, downstream receiver 34, upstream transmitter 28, queues data unit information database 30, segmenting and intermediate queue unit 29, reassembly unit 25 and a plurality of queues, such as PQ(1,1) – PQ(1,4) 31 – 37. Data unit information generator 21 may include a classifier, a policing unit and a marking unit. Data unit information generator 21 is coupled to packet distributor 19 for providing IP packets, and is coupled to data unit information database 30 for providing data unit information. The classifier analyzes the incoming (or reassembled) IP packets to determine to which class of service they belong. The policing unit is operable to enforce policy rules. For example, the policy rules can include a mechanism for dropping packets, such as but not limited to the WRED or RED mechanisms, in response to various parameters, including network traffic load, and the like. The policy rules may also determine when to change the class of service of incoming packets. The marking unit is operable to add information to the IP packets (to 'mark' them) in response to the classification and policing

NU 9 further includes a reassembly unit 25 that has an input 7 (second input of NU 9) for receiving fixed sized cells, such as ATM cells. Reassembly unit 25 is coupled to data unit information generator 21 for providing reconstructed IP packets. Reassembly unit 25 has a plurality of queues (denoted

VC/VP Q), each for a single combination of VP/VC fields of incoming ATM cells, such that it is able to reconstruct packets from ATM cells.

Exemplary segmentation and reassembly operations are illustrated in Figure 3A. A segmentation process starts by receiving IP packet 100 and ends by the provision of AAL5 compliant ATM cells 112 – 116. A reassembly process starts by receiving AAL5 compliant ATM cells 112 – 116 and ends by providing IP packet 100.

An IP packet 100 is converted to an AAL5 compliant packet 110 by adding to the IP packet 100 various fields such as IP type field 102, IP packet length field 106, CRC field 104 and stuffing bytes 108. IP packet length field 106 indicates the length of IP packet 100, and can be used as a data unit information unit. The length of stuffing bytes 108 is designed such that the overall length of AAL5 compliant packet 110 equals N * 48 bytes, N being a positive integer. As illustrated by figure 3A, the overall length of AAL5 compliant packet 110 is 3 * 48 bytes. AAL5 compliant packet 110 is converted to three AAL5 ATM cells 112, 114, and 116. These AAL5 ATM cells have the same VP/VC fields within their ATM headers. If is further noted that the same process can be applied to other protocols as well.

It is noted that the segmentation is not necessary according to another aspect of the invention.

Packet distributor 19 is coupled to a plurality of queues, such as PQ(1,1) – PQ(1,4) 31 – 37, and is operable to distribute the IP packets among the queues in response to the class of service of the IP packets.

PQ(1,1) – PQ(1,4) 31 – 37 are coupled to segmenting and intermediate queue unit 29, that is operable to segment IP packets to ATM cells, such as AAL5 compliant ATM cells, and to provide the ATM cells to upstream assembler and transmitter 28 . It is noted that in NU 8 of figure 1A IP packets are segmented before being stored in the data queues, while in NU 9 of figure 9 the IP packets are segmented only when they are scheduled to be upstream

transmitted. It is noted that the order of the storing and segmenting operations is not significant And In other implementation of the invention the segmentation operation (as well as segmenting and intermediate queue unit 29) can be omitted and native (variable sized) packets can be transmitted.

Segmenting and intermediate queue unit 29 is operable to segment the IP packets to ATM cells, and to store them in intermediate queues. In cases where not all the ATM cells of an ATM cell group that originated from a single IP packet can be transmitted, the remaining ATM cells, can be provided to upstream assembler and transmitter 28 during another timeslot. According to another aspect of the invention segments of packets are fetched from PQ(1,1) – PQ(1,4) 31 – 37 only if they can be transmitted and unit 29 does not store portions of packets.

Figure 1C illustrates network unit 9" that includes queues and scheduling and management entities that are able to handle both fixed sized cells and variable sized packets or frames. For convenience of explanation NU 9" is illustrated as a GFP standard compliant unit, but this is not necessarily so. Network unit 9" includes, in addition to various entities from NU 9 and NU 9' TDM over GFP unit 38" and packet distributor and optional framing unit 19", having framing capabilities that are required under the GFP protocol. In NU 9" data unit information may include both data relating to variable sized packets or frames and also information relating to groups of fixed sized cells.

Figure 1D illustrates network unit 9* that is operable to manage variable sized packets or frames without performing segmentation. Accordingly, the segmentation unit are omitted.

The Headend

Referring to Figure 2, illustrating Headend, such as OLT 38, in accordance with an embodiment of the invention.

OLT 38 includes (i) downstream data interface 40, (ii) grant controller 48, (iii) downstream transmitter 46, (iv) grant allocation unit 49, (v) upstream receiver 54, and (vi) grant queues GQ1 – GQ4 40 – 46

Downstream data interface 40 and grant allocation unit 49 are coupled to downstream transmitter 46. Grant controller 48 is coupled to upstream receiver 54 and to grant queues GQ1 – GQ4 40 – 46. Grant queues GQ1 – GQ4 40 – 46 are coupled to grant allocation unit 49.

Downstream data interface is operable to receive data to be downstream transmitted to NUs over a passive optical network and to provide the received data to downstream transmitter 46. Downstream transmitter 46 is operable to further receive data grants and data unit information request grants from grant allocation unit 48 and to generate downstream frames.

Upstream receiver 54 receives upstream-transmitted frames and is operable to extract data unit information out of the upstream-transmitted frames, to provide the data unit information to grant allocation unit 48 and to provide the upstream data along an upstream path from OLT 38 to other Headends, networks or other network units.

Grant controller 48 is operable to receive data unit information and to issue data grants in response to the data unit information. The allocated data grants are stored in grant queues GQ1 – GQ4 40 – 46, according to their class of service. Each grant queue GQ has a predefined class of service and holds all the grants for this class of service.

Grant queues GQ1 – GQ4 40 – 46 are coupled to grant allocation unit 49 that selectively fetches issued data grants according to a predefined order, and sends a sequence of data grants to downstream transmitter 46, where the data grants are assembled in to a downstream frame. For example, the selection may start by selecting an issued data grants of the highest class of service, and continue to the lower priority class of service grant queues, as long as additional timeslots may be allocated. It is noted that when implementing a fixed timeslot

allocation, such as in the case of type 1 T-CONTs, an additional unit, such as fixed allocation unit 41 may be operable to force grant allocation unit 49 to generate data grants in accordance with a fixed timeslot allocation.

The data unit information request algorithm

This algorithm is responsive to determine the data unit information to be sent from network units to the Headend. The algorithm evokes every cycle. In every cycle the Headend determines the capacity of data information to be received from each queue. The determination is followed by downstream transmissions of requests to receive data unit information from the network units. The network units transmit the data information according to the OLT request.

The algorithm assures that all the necessary information will be available to the scheduling algorithm of the MAC but that irrelevant information will not be transferred; so minimum bandwidth is used for the information transfer: This is implemented by: (a) transferring only the relevant information (information regarding data that might be transmitted in the next cycle), (b) only data unit information that was not transmitted previously is transferred, and (c) the data unit information content, rate and capacity are determined in accordance with the queue current status and the queue's class of service.

The algorithm distinguishes between different classes of service, and can handle each one differently. The differences among classes of service can be (a) in the data unit information content (for some classes of service, we can use less accurate information regarding each data unit information, while for others we might need more details), and/or (b) in the data information updates rate, and/or (c) in the capacity of updates allocated for each queue according to its status (eg. in some classes of service, it is possible that not all the relevant data unit information is transmitted, such as when the pace of updates is slower than that of the incoming traffic, in some classes of service when a queue is empty a

recovery time will be needed before utilizing full bandwidth capacity while in other the full capacity is always available).

The algorithm is responsive to some considerations, such as the (i) transfer method and, (ii) the upstream update rate. (i) Transfer method: the algorithm allocates update capacity for every queue in every time frame. The amount of data unit information that should be transferred is unknown when this size is calculated. On the one hand, this size must be large enough to include all possible input, but on the other hand, this size must be minimal so that bandwidth is not wasted (if for example this queue has no information to transfer). (ii) The upstream update rate: on one hand, high rate improves the delay and jitter performances, and the information received by the scheduler is more standard. But, on the other hand, in every update's transfer there is a bandwidth waste overhead, and as the rate is higher so is the percentage of the bandwidth that is wasted. Moreover, performances of the CPU must be taken into account, so that the execution time complies with the updates rate.

The following section further illustrates the algorithm.

These variables are defined:

Max Bytes In Queue – maximum bytes that can be sent from this queue if all bandwidth is allocated to it (this is limited by the total upstream rate). Max Bytes In Queue = Total Grants * N, whereas Total Grants - Number of grants to allocate in a time frame, and N – number of data bytes in an ATM cell (N = 48).

Actual Bytes In Queue – number of bytes in the mirror queue (A mirror queue is a representation of a queue within the data unit information database). This size updates constantly (when new packets are pushed into the queue and when packets are popped out when bandwidth is allocated for them).

Bytes In Update – Number of bytes that can be reported by the queue. (Update grants are given in accordance with this number).

Requested Bytes In Update – maximum number of bytes that might be in the ONT queue, but were not reported yet (this is the maximum relevant update size).

Max Input Bytes – maximum bytes that can enter the ONT queue in a time frame (this is limited by the input rate to this queue).

Actual Bytes In Update – the number of bytes that were received from a specific queue in this data unit information update.

Update Utilization – the percentage of utilization in the data unit information update (the ration between Actual Bytes In Update and Requested Bytes In Update).

Granted Bytes – the number of bytes that will be transmitted by a specific queue in this cycle (according to the data gmats that were allocated for this queue in the grant allocation)

The upstream update transfer method is described in the relevant section.

The algorithm:

(A) When the data unit information is received:

a. For every queue:

i. Extract all data unit information received for this queue.

ii. Calculate Update Utilization: Update Utilization =
(Actual Bytes In Update / Requested Bytes In Update).

iii. If Update Utilization < 100% (the last report was enough to report all the data in the queue) : Reset Requested Bytes In Update:
Requested Bytes In Update = (Max Input Bytes * RBIU_FACTOR(class of service id, Update_Utilization).
(When RBIU_Factor is defined here: For classes of service that have strict delay constrains (such as EF): RBIU_FACTOR = 1.
For other classes of service: RBIU_FACTOR = X. When X is determined
according to the class of service. The algorithm handles different Classes

of service differently: This is by allocating maximum update capacity for some classes of service while decreasing the update capacity below the maximum possible input for others).

iv. Push the data unit information to the relevant mirror queue.

(B) During Grant allocation:

a. For every queue:

i. Calculate Actual Bytes In Queue:

Actual Bytes In Queue = (Actual Bytes In Queue – Granted Bytes).

(C) Determining Data Unit information updates:

a. For every queue:

i. Calculate Bytes In Update; Bytes In Update = MIN (Requested Bytes In Update,

(Max Bytes In Queue - Actual Bytes In Queue).

(For every queue the Bytes in Update are in accordance with the amount of data that might be upstream transmitted in the next cycle). Note that Bytes In Update can also be controlled according to the class of service. The Bytes In Update can be decreased for lower classes of service when the total bandwidth is fully utilized. This is by re-computing Bytes In Update: Bytes In Update = Bytes In Update * BIU_Factor(class of service Id, Total Bandwidth allocation).

ii. Update Requested Bytes In Update: Requested Bytes In Update = (Requested Bytes In Update + (Max Input Bytes - Bytes In Update)).

iii. Allocate data unit information grants: #_Data_Unit_Information_Units = DUIU(class of service Id, Bytes In Update). Whereas DUIU is computed like this:

DIUU = Constant (= 2) for classes of service in which minimum data unit information is requires (eg. #of ATM cells in the queue).

DUIU = Bytes In Update / MIN_PACKET_SIZE; for classes of service in high priority.

DUIU = Bytes In Update / MEAN_PACKET_SIZE; for classes of service in low priority.

(This means that we enable high priority classes of service to report all their data unit information, and as a result waste some upstream bandwidth (since we didn't use mean packet size), where as low priority classes of service might not report all their data unit information (eg. When a burst of small packets arrives),

Note that MEAN_PACKET_SIZE can be calculated for each queue, and/or each class of service, and/or each NU.

It is yet further noted that there must be a mechanism of synchronization. This mechanism will synchronize the ONT queues and the mirrors queues once in several cycles.

An example of the amount of data unit information slots and other system parameters is illustrated below.

It is assumed that: (i) a MAC cycle includes three T-frame, (ii) downstream data rate of 622 Mbytes/sec, (iii) upstream data rate of 622 Mbytes/sec, (iv) an IP packet is represented by two to five ATM cells; (v) a network unit that is connected to 100baseT communication lines receives a maximal amount of thirty IP packets, each being 64 byte long, per MAC cycle, it typically receives ten IP packets, each being 200 byte long,(vi) a network unit that is connected to 10baseT communication lines receives a maximal amount of three IP packets, each being 64 byte long, per MAC cycle, it typically receives a single 200 bytes long IP packet, (vii) a plurality of network units share the same upstream channel, about four to eight of them are connected to 100baseT communication lines, the other network units are connected to 10baseT communication lines.

Accordingly, during a single MAC cycle a maximal amount of 318 IP packets are received, typically only 129 IP packets are received (Limited by the total upstream bandwidth). A data unit information unit is a byte long, but may also be two bytes long. Assuming that the transmission of data unit information requires additional bytes, for example, for re-transmitting data unit information,

for transmitting CRC bits and the like, then each MAC cycle an amount of 3*318*1 bytes, that can be encapsulated within 20 ATM cells, are required. Therefore only (20 / 636 =) 3.15% of the upstream bandwidth is wasted, while this percentage can be decreased when using the methods stated here.

The grant allocation unit

Reference is now made to Figure 4 which illustrates an exemplary portion of grant controller 48. Grant controller 48 includes data unit information database 50, a plurality of arbitration units such as arbitrators 61 – 64, and data converter 65.

Portion 51_1 of data unit information database 50 stores the data unit information of DQ(1,1) - a list of "net" sizes/length of relevant payload of seventy consecutive ATM cell groups that are stored at DQ(1,1): SZ(1,1,1) – SZ(1,1,70). SZ(1,1,1) being the size of the group of ATM cells that is located at the head of DQ(1,1). Portion 51_32 of data unit information database 50 stores the data unit information of DQ(32,1) - a list of "net" sizes of sixty consecutive ATM cell groups that are stored at DQ(32,1): SZ(1,32,1) – SZ(1,32,60). Portions 51_1 – 51_32 of queues database 50 that are associated with a first class of service are accessible to a first arbitrator 61 that determines which first class of service ATM cell group to transmit. Portion 52_1 of data unit information database 50 stores the data unit information of DQ(1,2) - a list of "net" sizes of fifty four consecutive ATM cell groups that are stored at DQ(1,2): SZ(2,1,1) – SZ(2,1,54). SZ(2,1,1) being the size of the group of ATM cells that is located at the head of DQ(1,2). Portion 52_32 stores the data unit information of DQ(32,2) - a list of "net" sizes of sixty three consecutive ATM cell groups that are stored at DQ(32,2): SZ(2,32,1) – SZ(2,32,63). Portions 52_1 – 52_32 of queues database 50 are associated with a second class of service are accessible to a second arbitrator 62 that determines which second class of service ATM cell group to transmit.

Portion 53_1 of data unit information database 50 stores the data unit information of DQ(1,3) - a list of "net" sizes of twenty five consecutive ATM cell groups that are stored at DQ(1,3): SZ(3,1,1) – SZ(3,1,25). SZ(3,1,1) being the size of the group of ATM cells that is located at the head of DQ(1,3). Portion 53_32 stores the data unit information of DQ(32,3) - a list of "net" sizes of fifteen consecutive ATM cell groups that are stored at DQ(32,3): SZ(3,32,1) – SZ(3,32,15). Portions 52_1 – 52_32 of queues database 50 are associated with a third class of service are accessible to a third arbitrator 63 that determines which third class of service ATM cell group to transmit.

Portion 54_1 of data unit information database 50 stores the data unit information of DQ(1,4) - a list of "net" sizes of sixty consecutive ATM cell groups that are stored at DQ(1,4): SZ(4,1,1) – SZ(4,1,60). SZ(4,1,1) being the size of the group of ATM cells that is located at the head of DQ(1,4). Portion 54_32 stores the data unit information of DQ(32,4) - a list of "net" sizes of ninety consecutive ATM cell groups that are stored at DQ(32,4): SZ(4,32,1) – SZ(4,32,90). Portions 54_1 – 54_32 of queues database 50 are associated with a fourth class of service are accessible to a fourth arbitrator 64 that determines which fourth class of service ATM cell group to transmit. The results of the arbitration cycles are fed to data converter 65 and then to grant queues GQ1 – GQ4 40 - 46.

Data converter 65 is operable to convert a result of an arbitration, that usually reflects the "net' size of ATM payload to the "gross" amount of data to be actually transmitted . The "net size is also referred to as ATM payload size, to an amount of timeslot to allocate to the transmission. Accordingly, GQ1 - GQ4 40 – 46 store the amount of consecutive timeslots to allocate for an upstream transmission from a certain NU queue. These amounts are also referred to as issued data grants.

According to another embodiment of the invention the sized reflect the packets (or frames) that are transmitted over the shared media and data converter 65 is omitted.

According to an aspect of the invention, OLT 38 determines the amount of data unit information to receive from each NU. This determination reduces the amount of upstream bandwidth assigned to the transmission of data unit information.

It is noted that during a MAC cycle a maximal predefined amount of data may be upstream transmitted from each queue. The maximal amount is limited by upstream transmission rates and by the rate in which the upstream data is provided to the queue. The maximal amount can be predefined or can be dynamically changed to reflect various changes in the network.

At each MAC cycle OLT 38 determines the data unit information to be upstream transmitted by the NUs. The amount of data unit information is responsive to the data unit information that has already been received by OLT 38, the maximal predetermined amount and an estimation of the ATM cell groups that are stored at the queues of which their data unit information was not yet transmitted to OLT 38. An estimation of ATM cell groups size that are supposed to be stored in a data queue can be based upon statistical analysis of ATM cell groups received from (i) that data queue, (ii) queues that belong to the same T-CONT (in case where a single T-CONT has more than a single queue), (iii) queues that belong to the same ONU, (iv) queues that service the same class of service, from ATM cell groups received from a certain application, from ATM cell groups that originated from packets of a certain networking protocol, from ATM groups originating from the same source and/or destined to the same destination, and the like.

For example, and referring to Figure 4, it is assumed that (a) the data unit information is the length (in bytes) of IP packets from which ATM cells were generated (b) the received data streams are Ethernet compliant and the minimal

size of an Ethernet packet is 64 bytes, (c) the maximal amount of bytes that can be transmitted from data queue DQ(32,4) during a single MAC cycle is 5000, (d) portion 54_32 stores SZ(4,32,1) – SZ(4,32,90), (e) the aggregate size of SZ(4,32,1) – SZ(4,32,90) is 4540. Then OLT 38 will request ONU to report the sizes of the next ten variable size packets within DQ(32,4) : (5000-4540)/64 = 10.

Data structures

Referring to figure 3B illustrating downstream frames 120 - 140, in accordance to embodiments of the invention. The upper part of Figure 3B illustrates three downstream frames, such as T-frames, that are transmitted during a single MAC cycle. Each T-frame out of T-frames 120, 130 and 140 includes two hundred and twenty four slots. T-frame 120 includes two hundred and sixteen ATM cell slots and eight framing slots, also referred to as PLOAM slots. A first PLOAM slot 121 is followed by twenty seven ATM slots 122 , second PLOAM slot 121 and twenty seven ATM slots 122 and so on.

T-frame 120 is followed by T-frame 130 that includes two hundred and sixteen ATM cell slots and eight framing slots, also referred to as PLOAM slots. A first PLOAM slot 131 is followed by twenty seven ATM slots 132, second PLOAM slot 131 and twenty seven ATM slots 132and so on.

T-frame 130 is followed by T-frame 140 that includes two hundred and eighteen ATM cell slots, eight framing (PLOAM) slots and six DUI (Note that all GI should be replaced by DUI) slots. A first PLOAM slot 141 is followed by twenty seven ATM slots 142, second PLOAM slot 141, twenty seven ATM slots 142, and so on. The eighth PLOAM slot is followed by twenty one ATM slots and six GI slots in which a data unit information request vector 146 is transmitted. It is noted that six ATM cells is the maximum needed, when referring to thirty-two NU with four queues in each NU and two bytes request for each queue,.

The mentioned above format is suited to an optical passive network in which the downstream data rate equals the data upstream rate. Accordingly, each group of PLOAM slots in a T-Frame determines which upstream data to transmit during an upstream frame of two hundred and twelve slots.

The mention above is only a non-limiting implementation of the invention The same invention may be applied mutatis mutandis using GRANT field of the PLOAM cells defined in G.983.1, or any other way that can implement the transmission of the data unit information request vector to the NUs.

Six DUI slots 144 are allocated for transmitting data unit information requests that determine which data unit information to transmit upstream during a MAC cycle of three T-frames.

The lower part of Figure 3B illustrates a MAC cycle that includes three T-frames, that is suited for an optical passive network in which the downstream data rate is four times faster than the upstream data rate. In such a case the data unit information request vector will be planted into the empty PLOAMS. It is noted that there are six empty PLOAMS in every T-frame, However there are only twenty seven bytes in each PLOAM (as apposed to 48 bytes in a normal ATM cell) therefore, twelve PLOAM are needed to transfer the information. It is noted that the third T-frame (140) can be utilized to convey more downstream ATM cells or can include more data unit information request slots.

Referring to Figure 5A illustrating data unit information request vectors 152, 153 and 154 in accordance to embodiments of the invention.

Data unit information request vector 153 includes two types of fields - a first type field that indicates an identity of a queue and a second type indicating the amount of data unit information, relating to that queue, to be transmitted. For example, field 153_1 "REQUEST TO REPORT DUI RELATING TO THE Y1 QUEUE" is followed by field 153_2 "REQUEST TO REPORT DUI OF X1 GROUPS". These two fields are interpreted as a request to upstream transmit

data unit information relating to X1 consecutive ATM cell groups that are stored at the Y1'th queue, starting from the first ATM cell group which its data unit information was not reported yet. Each NU monitors the data unit information of each of the NU queues and tracks the previously upstream transmitted data unit information. Accordingly, each NU is able to determine which data unit information to prepare for upstream transmission

Data unit information request vector 152 includes only the second type of fields and may be transmitted when the order of the relevant queues is predefined.

Data unit information request vector 154 is compressed according to a predefined compression scheme. It includes compressed fields of the second type only, although this is not necessarily so. For example, a compressed data unit information request vector may also include the two types of fields.

Data unit information request vector 153 includes a request to transmit data unit information relating to X1 ATM cell groups from queue Y1 (fields 153_1 and 153_2), a request to transmit data unit information relating to X2 ATM cell groups from queue Y2 (fields 153_3 and 153_4), and a request to transmit data unit information relating to X3 ATM cell groups from queue Y3 (fields 153_5 and 153_6).

Data unit information request vector 152 includes a request to transmit data unit information relating to X1 ATM cell groups from queue Y1 (field 152_1, the identity of the queue is predefined), a request to transmit data unit information relating to X2 ATM cell groups from queue Y2 (field 152_2, the identity of the queue is predefined), a request to transmit data unit information relating to X3 ATM cell groups from queue Y3 (fields 152_3, the identity of the queue is predefined) and a request to transmit data unit information relating to X4 ATM cell groups from queue Y4 (fields 152_4 , the identity of the queue is predefined).

Data unit information request vector 154 includes the same requests as data unit information request vector 152, but in a compressed from.

Reference is now made to Figure 5B, illustrating data unit information vectors 160 and 170 in accordance to embodiments of the invention.

Data unit information vector 170 is a compressed version of data unit information vector 160.

The timing of the upstream transmission of each portion of data unit information vectors is determined by OLT 38 and may either be predefined or included within data unit information request vectors. (eg. If the order is predefined and all the data unit information is transmitted in one piece, one NU after the other, than Each NU can calculate its exact transmission time according to its predefined position and the sizes of the previous NUs data information reports. This update vector length can be constant or variable. In the latter case, the OLT only has to transmit the start time of this update vector, and each NU should recalculate its exact transmission time every cycle.).

Data unit information vectors 160 and 170 are transmitted in response to a reception of DUI request vectors. As mentioned above, the DUI vector may be transmitted in various manners. Accordingly the DIU vectors may be transmitted in response to (i) a reception of fields 152_1 and 152_2 of data unit information request vector 152, or (ii) a reception of fields 154_1 and 154_2 of data unit information request vector 154, or (iii) a reception of fields 153_1, 153_2, 153_3 and 153_4 of data unit information request vector 153 . Accordingly, the data unit information vector 160 starts by X1 fields 162_1 - 162_X1 that embeds the data unit information relating to X1 ATM cell groups being stored within queue Y1. These fields are followed by X2 fields 163_1 – 163_X2 that embed the data unit information relating to X2 ATM cell groups being stored within queue Y2.

Data unit information vector 170 includes X1 fields 172_1 – 172_X1 that embed compressed data unit information relating to X1 ATM cell groups being

stored within queue Y1. These fields are followed by X2 fields 173_1 – 173_X2 that embed compressed data unit information relating to X2 ATM cell groups being stored within queue Y2.

It is noted that the timing of the upstream transmission may be predetermined, negotiated between the OLT 38 and NUs or be included within the data unit information request vector Figure 5C illustrates an upstream frame, in accordance with an embodiment of the invention.

The upstream frame includes three upstream, such as T-frames 180, 190 and 200. Each T-frame includes 53 slots.

In some cases a single vector is located at the end of the three T-frames . This vector includes a group of ATM cells, when each NU transmits in only part of the ATM cell, and that the vector usually includes a CRC and Header.

Portions of data unit information vectors (also referred to as DUIV), are transmitting in predefined slots, such as DUIVs 210_1, 210_2 and 210_3 OLT 38 determines the position (timing) of each of these slots and also determines which data unit information to send during each slot. Each T-frame out of T-frames 180, 190 and 200 starts by a DUI vector slot, but this is not necessarily so . The inclusion of a DUI vector slot within each T-frame allows to update the GI of queues several times during a MAC cycle, thus allowing relatively frequent update of the timing information, and especially allows for upstream transmitting data unit information of high priority (high class of service), thus allowing to issue upstream transmit the high priority information before the current MAC cycle ends. According to another aspect of the invention the DUIV is transmitted in the end of the cycle.

The length of each DUIV slot may be predefined but it also may be configured to match the length of data unit information vectors from a certain set of queues. For example, the length (in ATM cells) of 210_1, 210_2 and 210_3 are (Z1-1), (Z2-1) and (Z3-1) accordingly.

<u>The arbitration mechanisms</u>

Reference is now made to Figures 6A – 6C, illustrating various arbitration mechanisms, in accordance with preferred embodiments of the invention. These arbitration schemes may be implemented by at least one of arbitrator out of arbitrators 61 – 64.

Conveniently, each arbitrator is operable to allocate resources/ grant upstream data transmission of data of a certain service of class. Arbitrator 61 is operable to grant upstream data transmissions of a first class of service, arbitrator 62 is operable to grant upstream data transmissions of a second class of service, arbitrator 63 is operable to grant upstream data transmissions of a third class of service, arbitrator 64 is operable to grant upstream data transmissions of the fourth class of service. Each of the class of service is associated with a distinct priority level. It is noted that at many networks the highest priority class of service receives a fixed bandwidth allocation, regardless of amount of traffic for that class of service. It is noted that the result of an arbitration process, such as an arbitration process implemented by arbitrators 61 – 64, is a data grant, The data grants are further selectively fetched by grant allocation unit 49, that actually allocates upstream timeslots. It is noted that the result may also be a responsive to the size of payload to transmit and this payload size is further translated to an overall size (including the size of additional headers, padding bytes and the like). Such a translation may be implemented by data converter 65 according to one embodiment of the invention but may also be unnecessary according to another aspect

Figures 6A - 6C illustrate arbitration mechanisms that are adapted to the transmission of cells across a network. According to another aspect of the invention packets and not cells are transmitted over the passive optical network. Accordingly the segmentation to cells and the overhead associated with the segmentation is not relevant, and the arbitration scheme is responsive to the size of the packets to be transmitted.

Figure 6A is a flow chart illustrating a first arbitration mechanism 310. First arbitration mechanism 310 is operable to allocate "fixed" data grants - data grants for transmitting a fixed amount of data. First arbitration mechanism 310 starts by a step 311 of defining a fixed amount of slot quota to allocate to each queue out of a set of queues that are arbitrated by first arbitration process 310. Usually, the set include queues of the same class of service. It is noted that queues that are associated to the same class of service may be arbitrated by a single arbitration process, but this is not necessarily so and they may be partitioned to a plurality of sets, each set of a predefined priority. It is further noted that each queue may be assigned a different fixed amount of slot quota.

Step 311 is followed by step 312 of allocating data grants in accordance with definitions of step 311. The data grant allocation is usually done regardless the emptiness of the queues. Step 312 may be executed each MAC cycle.

Figure 6B is a flow chart illustrating a second arbitration mechanism 320. Second arbitration mechanism 320 is operable to allocate a predefined amount of data grants in response to an indication that a queue that takes part in the arbitration cycle is not empty. Such a queue is referred to as a participating queue.

Second arbitration mechanism 320 starts by a step 321 of defining an amount of slot quota to allocate to each participating queue during an arbitration cycle. Step 321 is followed by step 322 of receiving an indication of the participating queues (which queues out of possible participating queues are not empty). This indication may be derived from previously received data unit information. Step 322 is followed by a step 324 of allocating data grants to participating queues.

Figure 6C is a flow chart illustrating a third arbitration mechanism 330. Third arbitration mechanism 330 is operable to allocate data grants in response to a predefined credit assigned to each queue and to the data unit information.

Third arbitration mechanism 330 starts by a step 332 of defining a "credit" quote for each queue that is arbitrated by third arbitration mechanism 330. The

"credit" quote is an amount of bytes that can be allocated to upstream data transmission from the queue. The credit quote assigned to each queue may reflect the queue priority / relative allocation of data grants. According to an aspect of the invention that the credit quota and/or can be dynamically changed . Step 332 is followed by step 334 of receiving data unit information relating to queues that are arbitrated by third arbitration mechanism 330.

Step 334 is followed by step 336 of determining from which queue to start the current arbitration cycle. Conveniently, the queues are arranged in a cyclical manner and an arbitration cycle starts from a queue that follows the last queue that was processed during a previous arbitration cycle.

Step 336 is followed by step 338 of determining whether the current queue is empty. If the answer is "yes" step 338 is followed by step 340 of selecting a new current queue if the current arbitration cycle did not end. If the arbitration cycle did not end step 340 is followed by step 338. If the arbitration cycle ended step 340 is followed by step 350 of ending the arbitration cycle. If the queues are arranged in a round robin formation, the next current queue is an adjacent queue of the round robin formation. If the queue is not empty, step 338 is followed by step 342.

Step 342 includes determining the amount of bytes that can be transmitted given the current credit of the current queue. It is noted that during a first iteration of the arbitration mechanism the current credit may equal the credit quote, but as further illustrated, the current quote is updated during the arbitration process. The ATM cell group may include consecutive ATM cells, starting with the first ATM cell group which is the ATM cell that is stored at the head of the current queue.

According to an aspect of the invention step 342 includes determining a "net" amount of data that may be transmitted, whereas the "net" amount is further converted to an amount of data cells. This conversion may be omitted if

variable sized packets or frames are transmitted over the shared media without being segmented.

If the current credit does not allow for transmitting at least the first data unit (for example, the first ATM cell group then step 342 is followed by step 343 of increasing the current credit by the credit quota. Step 343 is followed by step 340.

If the current credit does allow for transmitting a set of consecutive ATM groups, step 342 is followed by step 344 of allocating data grants for the transmission of the set and updating the current credit by adding a credit quota and decreasing the aggregate size of the ATM cell groups that form the set. Step 344 is followed by step 340.

Step 344 is followed by step 346 of determining when the arbitration cycle ends. The determination can be responsive to the amount of data grants that was allocated or to the amount of iterations of steps 338-342 or when all the queues where handled. The overall amount of allowed data grants can be dynamically changed.

If the arbitration cycle ended, step 346 is followed by step 350 ending the arbitration cycle. Else, step 346 is followed by step 340 of selecting a new current queue.

Distinct arbitration schemes, such as but not limited to arbitration schemes 310 – 330, may be implemented by various arbitrators, thus allowing for tailoring arbitration schemes to the characteristics of each class of service. These characteristics may include, bursty behavior patterns, delay sensitivity and the like.

Systems and methods that are adapted to handle various classes of service include dependent arbitrators and independent arbitrators. Independent arbitrators allocate data grants regardless the allocation of data grants by other arbitrators. Dependent arbitrators allocate data grants in response to data grant allocated by other arbitrators. Independent arbitrators are usually utilized for the

highest (most important) class of services and to fixed arbitration schemes. Dependent arbitrators are usually utilized for lower class of services. Dependent arbitrators may receive various input from other arbitrators or from the grant allocation unit), such as a residual overall data grant amount to be allocated during an arbitration cycle, or a relationship between allocated data grants to queues. In dependent arbitrators that implement the third arbitration mechanism 330 the credit quota, the current credit and the total amount of data grant location per arbitration cycle can be dynamically changed in response to the receive input from other arbitrators.

For example, systems and methods that are operable to handle the following class of services (provisioned bandwidth, guaranteed bandwidth, assured bandwidth and non-assured bandwidth) may (i) utilize an independent arbitrator that implements a first arbitration mechanism 310 for handling provisioned bandwidth, (ii) utilize a dependent arbitrator that depends upon the implements a second arbitration mechanism 320 for handling allocated bandwidth, the dependent arbitrators receives the amount of overall allowed data grants, given the results of the independent arbitrator , (iii) utilize a dependent arbitrator that implements a third arbitration mechanism 330 for handling provisioned bandwidth , the dependent arbitrators receives the amount of overall allowed data grants, given the results of the independent arbitrator and the mentioned above dependent arbitrator. The dependent arbitrator can also receive a relationship between data grant allocations of distinct queues in response to the allocation of the previous arbitrator, and (iv) utilize a dependent arbitrator that implements a third arbitration mechanism 330 for handling non-assured bandwidth. It is noted that other combinations of arbitration schemes may be implemented to fit various class of services, such as but not limited to provisioned bandwidth, minimum latency, assured bandwidth, non-assured bandwidth, minimum drop, minimum jitter and a combination of the mentioned above class of services.

It is noted that other configurations of arbitrations units may be implemented, such as multiple independent arbitrators and a single dependent arbitrator, or vice verse, but this is not necessarily so

An exemplary arbitration and grant allocation algorithm

The following exemplary scheduling algorithm is applicable to schedule the transmission of data units that may be associated with the following class of services: provisioned bandwidth (highest priority), guaranteed bandwidth, assured bandwidth, and non-assured bandwidth (lowest priority).

It is noted that this algorithm evokes every cycle. The number of grants to allocate in every cycle (Total Grants) = T * R * X. Whereas T – number of ATM cells in one T-frame for an upstream rate of 155 Mbit/sec (T = 53). R – ratio between upstream rate and 155 Mbit/sec. X – number of T-frames in a cycle. And T * R – number of ATM cells in one T-frame for the current upstream rate.

The algorithm schedules a transmission of data units in response to a data unit such as data information unit data base 50 of Figure 2.

The algorithm starts by few arbitration sections. The arbitration sections are followed by grant allocation steps.

(i) Provisioned bandwidth arbitration section. Prior to this section an index is assigned to each provisioned bandwidth queue.

a. If this grant should be allocated to this queue in this time frame (as provisioned) – this grant will be allocated to this queue (whether a grant should be allocated or not is a question of this T-cont rate. Every T-cont should get a grant once every Y time frames to comply with the requested rate).

b. Otherwise – this grant will not be allocated in this section.

It is noted that when necessary (there are more than Total Grants Provisioned bandwidth T-CONTs) it is possible that one index will be assigned to several provisioned queues. If necessary (the rate of a Provisioned Bandwidth T-cont demands) it is possible that several indexes will be assigned to one provisioned queue. Unlike all other classes of Service – There is no FIFO for

Provisioned grants. In the Arbitration procedure these grants are planted into the designated index. Accordingly, this stage may be performed only during the arbitration procedure - when the grants are planted into the downstream PLOAMS.

(ii) Guaranteed bandwidth arbitration section. During this section performing Round Robin algorithm among all Guaranteed bandwidth queues. The Round Robin algorithm includes the steps of: Start from one of the queues. Go over all queues, until all queues are empty: If there is data waiting in the queue – allocate the equivalent number of ATM cells for that queue so that all data can be sent (mandatory information required – number of ATM cells needed for sending all pending data), and push the allocated grants into Guaranteed bandwidth grants FIFO queue.

It should be noted that this implementation premises that this class of service is policed beforehand (i.e. there is no excessive use). If this premise is not supported than excessive use must be prevented (by using leaky bucket mechanism or DRR). Information regarding the exact net data capacity is then needed in order to perform this. Note that usually, in this kind of traffic, when delay and jitter performances are highly important, excessive data is discarded.

It is further noted that additional improvements can be done in the this part of the scheduling mechanism: If more information is available (i.e. internal packets division) than only one packet can be sent in each round from each queue. This ensures greater fairness. If needed – hierarchy can be set among different Classes of service from this type, where the scheduling will be performed one after the other. In this case – there must be one FIFO for every hierarchy.

(iii)Assured bandwidth arbitration section. This section includes performing Deficit Round Robin (DRR) algorithm among all Assured bandwidth queues. A DRR includes the steps of : Start from the first queue. Go over all queues (one round exactly). Perform DRR based on: packets size and credit. In each queue's

turn, bandwidth is allocated for all packets that can be sent according to the credit. The DRR is followed by allocating the equivalent number of ATM cells for the packets that were scheduled and pushing the allocated grants into Guaranteed bandwidth grants FIFO queues.

Note the due to the nature of the reports described in the invention a DRR or other packet scheduling mechanism can be implemented since the arbitrators has information about packets (like packet length) and not just the queue length.

It is noted that a hierarchy can be set among different Classes of service from this same basic Service Class, the scheduling will be performed one after the other. In this case – there must be one FIFO for every hierarchy.

(iv) Non-Assured and Best Effort bandwidth arbitration section. This section is executed only if there are still some remaining data grants to allocate after the previous sections were executed. This section includes performing DRR algorithm among all Non-Assured bandwidth queues, in a manner that resembles the allocation of assured bandwidth. It is noted that a hierarchy can be set among different Classes of service from this basic Service Class, where the scheduling will be performed one after the other. In this case – there must be one FIFO for every hierarchy.

(v) Surplus bandwidth arbitration section: if there are still any data grants to allocate, an additional allocation mat take place. During this section data grants are allocated for transmitting data units of other classes of service (such as excess bandwidth for the guaranteed bandwidth).

It is noted that if needed a hierarchy can be set among different Classes of service from this basic Service Class, where the scheduling will be performed one after the other. In this case – there must be one FIFO for every hierarchy.

The arbitration steps are followed by grant allocation steps. It is noted that the result of the grant allocation steps is a data grant vector and that each queue is associated with an index.

(i) Provisioned Bandwidth grant allocation: "plant" data grants according to provisioned bandwidth arbitration scheme within the data grant vector. It is noted that provisioned bandwidth schemes determine both the length of allocated bandwidth and the location within each upstream frame.

(ii) Guaranteed Bandwidth grant allocation: retrieve data grants from the Guaranteed bandwidth grants queue until the Guaranteed bandwidth queue is empty or a predefined amount of data grants (Total Grants) are allocated. It is noted that if there is hierarchy between several classes of service from this allocation is repeated for each hierarchy order, in descending order.

(iii) Assured Bandwidth grant allocation. This allocation is done only if additional data grants may still be allocated after the previous data allocations are executed. This allocation includes retrieving data grants from the assured bandwidth grants queue until the assured bandwidth queue is empty or a predefined amount of data grants (Total Grants) are allocated. It is noted that if there is hierarchy between several classes of service from this allocation is repeated for each hierarchy order, in descending order.

(iv) Non -Assured Bandwidth and Best Effort grant allocation. This allocation is done only if additional data grants may still be allocated after the previous data allocations are executed. This allocation includes retrieving data grants from the non - assured bandwidth grants queue until the non-assured bandwidth queue is empty or a predefined amount of data grants (Total Grants) are allocated. It is noted that if there is hierarchy between several classes of service from this allocation is repeated for each hierarchy order, in descending order.

It is noted that since this is a dependent arbitrator – its arbitration algorithm can only take place after the allocation of the previous classes of service (since the amount of grants to allocate depend on the previous allocations), a delay is expected at this point.

(iv) Surplus bandwidth grant allocation: This allocation is done only if additional data grants may still be allocated after the previous data allocations are executed.

It is noted that In order to improve the delay and jitter performances of the Guaranteed bandwidth class of service , grants can be allocated immediately upon receiving these queues updates (without the delay that derives from the periodic execution of the algorithm). In this case the Guaranteed bandwidth reports will be handled immediately upon receiving and the relevant grants will be pushed into the relevant queue.

It is further noted that if the SLA addresses also ONT total bandwidth (that can be shared among all Classes of service), than these changes should be performed: (i) In the Assured bandwidth arbitration section, if a queue has credit and no Assured bandwidth data to send, packets from this ONT other queues should be sent (credit is changed accordingly).

If the SLA addresses also ISP total bandwidth (regardless of internal division) than these changes should be performed: All ISP's queues from the same Service Class will be stored in one queue in the OLT. In this way, allocation is according to ISP SLA. The information regarding each packet should also include source T-Cont.

The Deficit Round Robin algorithm includes the steps of:

a. Choosing a quantum of bits to serve from each connection in order.

b. Each connection has a deficit counter (to store credits) with initial value zero.

c. For each Head Of Line packet: If its size <= (quantum + credit) send and save excess. Otherwise save entire quantum. If no packet to send, reset counter.

It is noted that the assured bandwidth DRR implementation in this algorithm includes setting the Quantum to : Quantum = (Assured Bandwidth / Total Bandwidth) * (Total Grants * N). Whereas N – number of data bytes in an ATM cell (N = 48). (Total Grants * N) – total number of bytes in one cycle, and

(Assured Bandwidth / Total Bandwidth) – ratio of this connection in comparison with other connections.

It is noted that during every cycle (X T-Frames) exactly one round of DRR is performed. Therefore, It is easy to see that exactly the assured bandwidth is allocated (assuming the relevant queue is full).

It is noted that non-assured bandwidth DRR implementation in this algorithm includes setting the Quantum to Quantum = partial weight among all other queues.

Since this DRR is performed only when bandwidth is sufficient, the next queue to be serves must be saved, such that on every round the DRR distribution begins in this queue.

Also, it is possible that only part of a packet will be sent (if there are not enough grants for the whole packet). In this case the details for the packet that was partially sent must be updated (to reflect its current size), and this queue should be the first to be serves in the next round (to complete its turn).

It is noted that the method and apparatus according to the present invention can be implemented either in hardware, in software or in a combination thereof.

It will be apparent to those skilled in the art that the disclosed subject matter may be modified in numerous ways and may assume many embodiments other then the preferred form specifically set out and described above.

Accordingly, the above disclosed subject matter is to be considered illustrative and not restrictive, and to the maximum extent allowed by law, it is intended by the appended claims to cover all such modifications and other embodiments, which fall within the true spirit and scope of the present invention.

The scope of the invention is to be determined by the broadest permissible interpretation of the following claims and their equivalents rather then the foregoing detailed description.

WE CLAIM

1. A communication system comprising: an optical communication network interconnecting a Headend and a plurality of network units; wherein the Headend has a media access controller for issuing data grants and data unit information requests; wherein a data grant being issued at least partially in response to previously received data unit information; and wherein at least some network units out of the plurality of network units are operable to: receive data to be transmitted to the Headend; transmit data unit information associated with the received data; and transmit data to the Headend in response to data grants issued by the media access controller.

2. The system according to claim 1 wherein a sequence of data grant authorizes an identified network unit out of the plurality of network units to transmit a group of consecutive data cells during consecutive timeslots.

3. The system according to claim 1 wherein the media access controller comprises a plurality of arbitrators.

4. The system according to claim 3 wherein each arbitrator is associated with a class of service out of a plurality of class of services.

5. The system according to claim 4 wherein at least one class of service is characterized by at lest one of the following parameters: provisioned bandwidth, guaranteed bandwidth, assured bandwidth, non-assured bandwidth or surplus bandwidth.

6. The system according to claim 4 wherein at least one class of service is characterized by at least one of the following parameters: minimum latency, , minimum bandwidth, maximum bandwidth, maximum burst size, minimum drop, and minimum jitter.

7. The system according to claim 4 wherein at least one class of service is a Diffserv class of service.

8. The system according to claim 4 wherein at least one class of service is an IntServ class of service.

9. The system according to claim 4 wherein at least one class of service is an MPLS class of service.

10. The system according to claim 1 wherein the Headend is operable to transmit data to network units in consecutive frames, wherein each frame comprises at least one data grant and at least one data unit information request.

11. The system according to claim 1 wherein data unit information represents at least one parameter of a group of data cells to be transmitted from the network unit.

12. The system according to claim 11 wherein each group of data cells comprises relevant payload and overhead signals.

13. The system according to claim 12 wherein data unit information reflects the length of the relevant payload.

14. The system according to claim 12 wherein data unit information reflects a time of arrival of the relevant payload.

15. The system according to claim 11 wherein data unit information reflects a length of at least one data unit to be transmitted from the network unit.

16. The system according to claim 11 wherein data unit information reflects a time of arrival of at least one data unit to be transmitted from the network unit.

17. The system according to claim 1 wherein data unit information represents at least one parameter of a variable length packet or frame to be transmitted from the network unit.

18. The system according to claim 17 wherein data unit information reflects the length of the variable length packet or frame.

19. The system according to claim 17 wherein data unit information reflects a time of arrival of the variable length packet or frame.

20. The system according to claim 1 wherein the media access controller is operable to determine an amount of data unit information to be sent from a network unit.

21. The system according to claim 20 wherein the determination is responsive to an estimation of data unit information to be sent from the network units.

22. The system according to claim 20 wherein the determination is responsive to data unit information previously transmitted from the network unit and to a data threshold.

23. The system according to claim 22 wherein the data threshold reflects a maximal amount of data that can be transmitted from the network unit to the Headend during a predefined time period.

24. The system according to claim 1 wherein at least some of the network units are not operable to generate data unit information; wherein the media access controller estimates data unit information relating to data being received from network units.

25. The system of claim 24 wherein the at least some network units further comprise a classifier, for classifying incoming data packets in response to their service of class.

26. The system of claim 1 wherein at least some of the network units are operable to generate data unit information reflecting either variable sized packets or frames or fixed length cells.

27. The system of claim 26 wherein the variable length data packets are Internet Protocol packets and the fixed sized cells are Asynchronous Transfer Mode cells.

28. The system of claim 26 wherein the variable length data frames are Ethernet frames and the fixed sized cells are Asynchronous Transfer Mode cells.

29. The system according to claim 1 wherein the media access controller is operable to issue data grants in response to at least one arbitration scheme.

30. The system of claim 29 wherein the received data belongs to a class of service out of a plurality of class of services.

31. The system of claim 29 wherein one arbitration scheme allocates data grants in a fixed manner.

32. The system of claim 29 wherein one arbitration scheme allocates data grants in response to data unit information.

33. The system of claim 29 wherein one arbitration scheme allocates data grants in response to a transmission current credit.

34. A media access controller for controlling an access of a plurality of network units to a shared upstream channel, the media access controller being coupled to a receiver, for receiving data unit information from the plurality of network units; wherein the media access controller comprising at least one arbitration unit, coupled between the receiver and a grant allocation unit, for arbitrating between requests to upstream transmit data units.

35. The media access controller of claim 34 wherein the data units are groups of fixed sized cells and wherein the data unit information reflects at least one parameter of said groups.

36. The media access controller of claim 34 wherein the data units are variable sized packets or frames and wherein the data unit information reflects at least one parameter of said variable sized packets or frames.

37. The media access controller of claim 34 further comprising a grant allocation unit, for selecting data grants authorizing an upstream transmission of a group of fixed sized cells in response to the arbitration.

38. The media access controller according to claim 34 wherein fixed sized cell groups belong to a class of service out of a plurality of class of services; and wherein each arbitration unit out of the at least one arbitration units is operable to arbitrate between requests of at least the same class of service.

39. The media access controller according to claim 34 operable to implement at least one arbitration scheme, whereas at least one of said arbitration schemes includes allocating data grants in a fixed manner.

40. The media access controller according to claim 34 wherein data unit information reflects a length of a group of fixed sized cells or a length of a relevant payload embedded within said group.

41. The media access controller according to claim 34 wherein data unit information reflect a length of at least one data unit to be transmitted from the network unit.

42. The media access controller according to claim 34 wherein data unit information reflects a time of arrival of at least one data unit to be transmitted from the network unit.

43. The media access controller according to claim 34 wherein the grant allocation unit is operative to receive allocated data grants from the at least one arbitrating unit and to select data grants in response to a predefined priority between the at least one arbitration unit.

44. The media access controller according to claim 34 wherein the grant allocation unit is operative to receive allocated data grants from the at least one arbitrating unit and to select data grants in response to a predefined priority between classes of service to which the data grants are associated with.

45. The media access controller of claim 34 wherein at least one arbitration scheme is responsive to data unit information.

46. The media access controller of claim 34 wherein at least one arbitration scheme is responsive to a transmission current credit.

47. The media access controller of claim 34 further operable to determine an amount of data unit information to be sent from a network unit.

48. The media access controller of claim 47 wherein the determination is responsive to an estimation of data unit information to be sent from the network units.

49. The media access controller of claim 47 wherein the determination is responsive to data unit information previously transmitted from the network unit and to a data threshold.

50. A method for allocating upstream bandwidth of a shared upstream channel of an optical network, the optical network interconnecting a Headend with a plurality of network units, the method comprising the steps of: determining data unit information to be upstream transmitted from at least one network unit; receiving upstream transmitted data unit information; and issuing data grants authorizing an identified network unit to transmit upstream data in response to previously received data unit information.

51. The method according to claim 50 wherein the optical network is a passive optical network.

52. The method according to claim 50 wherein the step of issuing comprising the steps of: arbitrating between requests to transmit groups of fixed sized data cells; allocating data grants in response to the arbitrating; and selecting allocated data grants.

53. The method according to claim 50 wherein the step of arbitrating comprising performing at least two arbitration cycles; and wherein the step of selecting is responsive to a predefined priorities assigned to arbitration cycles.

54. The method according to claim 50 wherein the step of issuing comprising the steps of: arbitrating between requests to transmit a variable sized packet or

frame; allocating data grants in response to the arbitrating; and selecting allocated data grants.

55. The method of claim 50 wherein a variable sized packet is converted to a plurality of fixed sized cells, and wherein the data unit information reflects the total length of said plurality of fixed sized cells.

56. The method of claim 55 wherein the variable sized packet is an Internet Protocol packet and the fixed sized cell is an ATM cell.

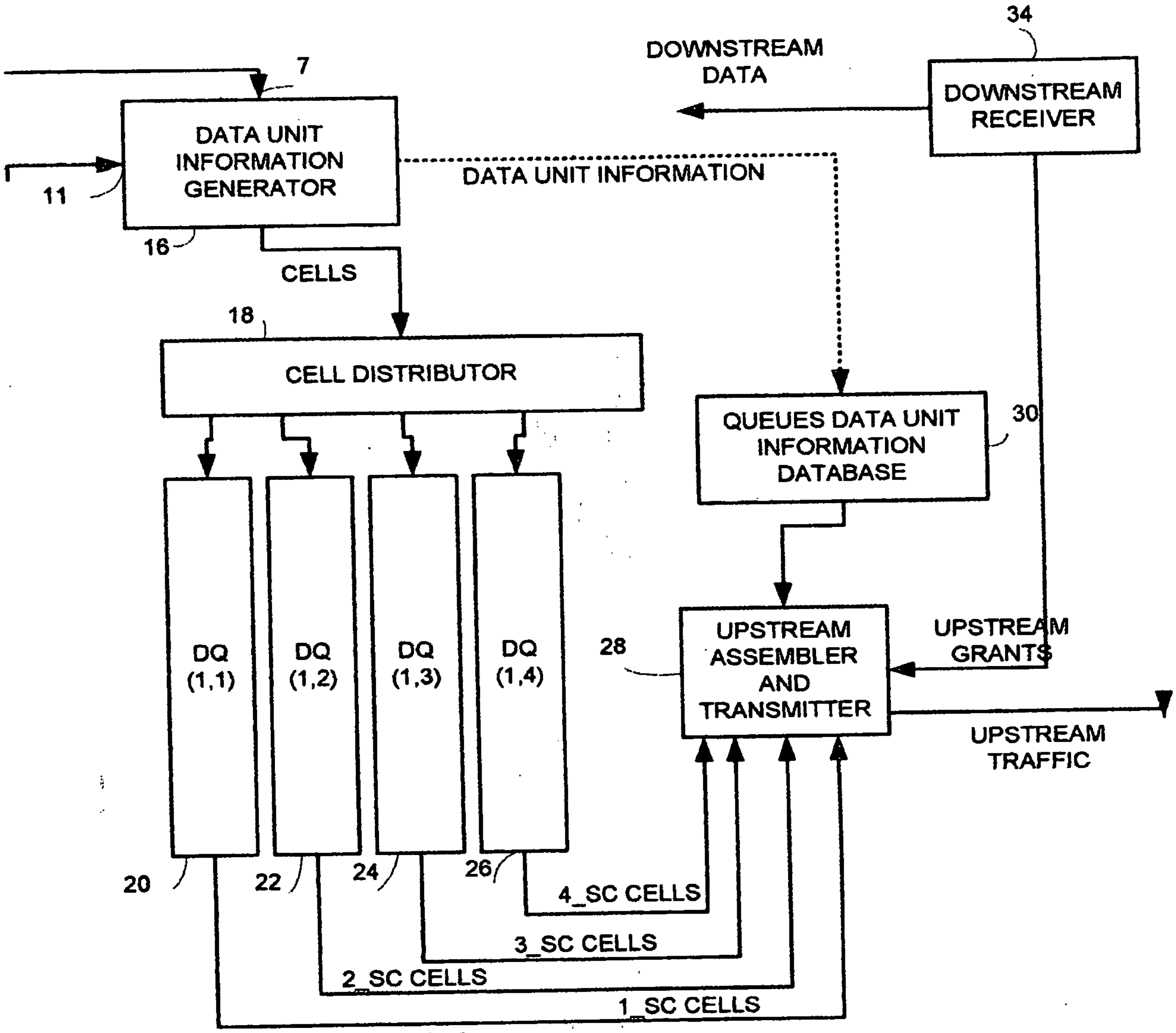
57. The method of claim 55 wherein the variable length data packets are Ethernet frames.

58. The method according to claim 50 further comprising a step of determining an amount of data unit information to be sent from a network unit.

59. The method according to claim 58 wherein the determination is responsive to an estimation of data unit information to be sent from the network units.

60. The method according to claim 58 wherein the determination is responsive to data unit information previously transmitted from the network unit and to a data threshold.

For the Applicants
Oren Reches, Adv., Patent Attorney



7
11
DATA UNIT INFORMATION GENERATOR
16
DOWNSTREAM DATA
34
DOWNSTREAM RECEIVER
DATA UNIT INFORMATION
CELLS
18
CELL DISTRIBUTOR
QUEUES DATA UNIT INFORMATION DATABASE
30
DQ (1,1)
DQ (1,2)
DQ (1,3)
DQ (1,4)
28
UPSTREAM ASSEMBLER AND TRANSMITTER
UPSTREAM GRANTS
UPSTREAM TRAFFIC
20
22
24
26
4_SC CELLS
3_SC CELLS
2_SC CELLS
1_SC CELLS

FIG. 1A

8

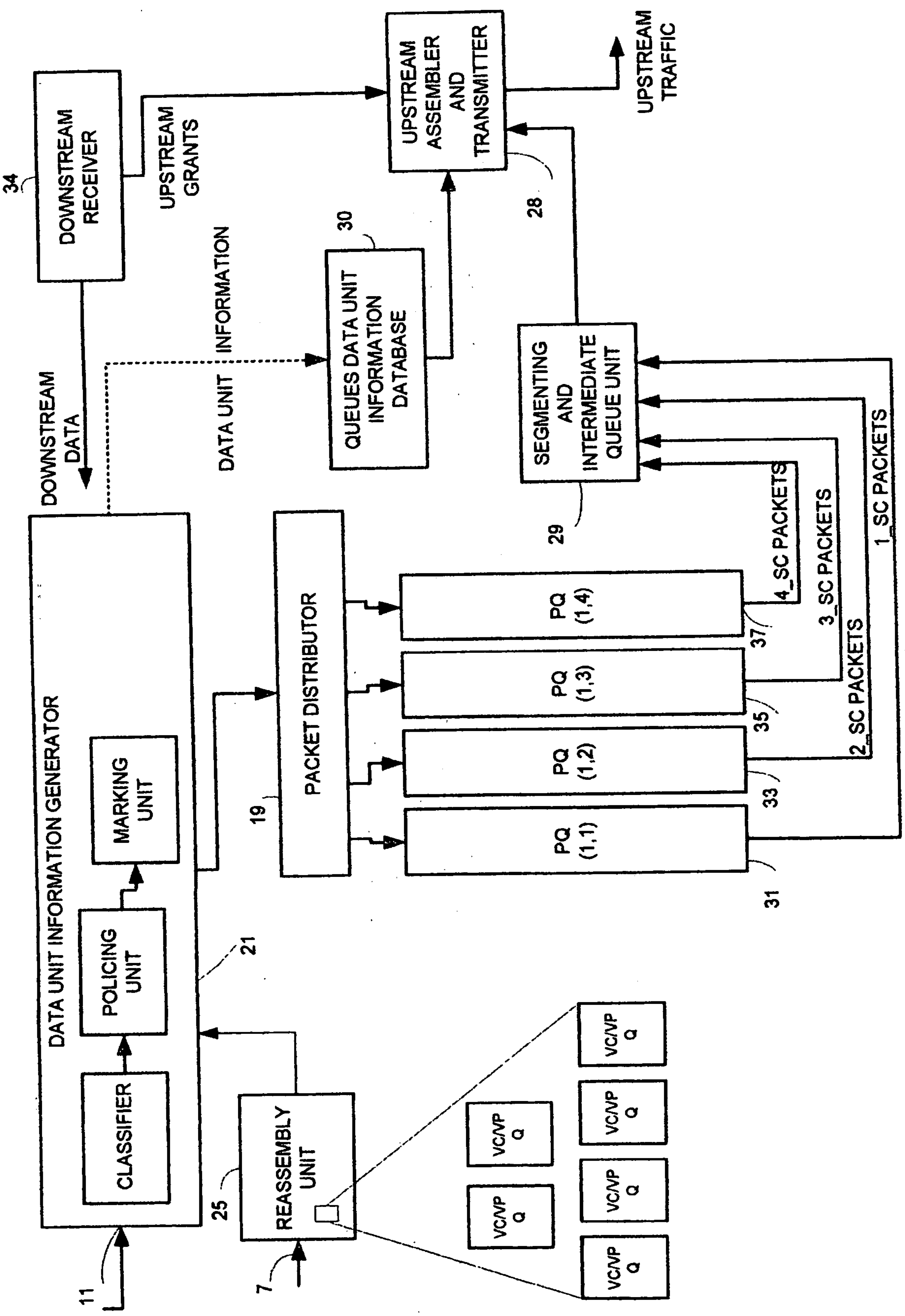

DOWNSTREAM RECEIVER
34
UPSTREAM GRANTS
UPSTREAM ASSEMBLER AND TRANSMITTER
UPSTREAM TRAFFIC
28
30
QUEUES DATA UNIT INFORMATION DATABASE
DATA UNIT INFORMATION
DOWNSTREAM DATA
SEGMENTING AND INTERMEDIATE QUEUE UNIT
29
DATA UNIT INFORMATION GENERATOR
MARKING UNIT
POLICING UNIT
CLASSIFIER
21
PACKET DISTRIBUTOR
19
PQ (1,4)
PQ (1,3)
PQ (1,2)
PQ (1,1)
37
35
33
31
4_SC PACKETS
3_SC PACKETS
2_SC PACKETS
1_SC PACKETS
REASSEMBLY UNIT
25
11
7
VC/VP Q
VC/VP Q
VC/VP Q
VC/VP Q
VC/VP Q
VC/VP Q

FIG. 1B 9

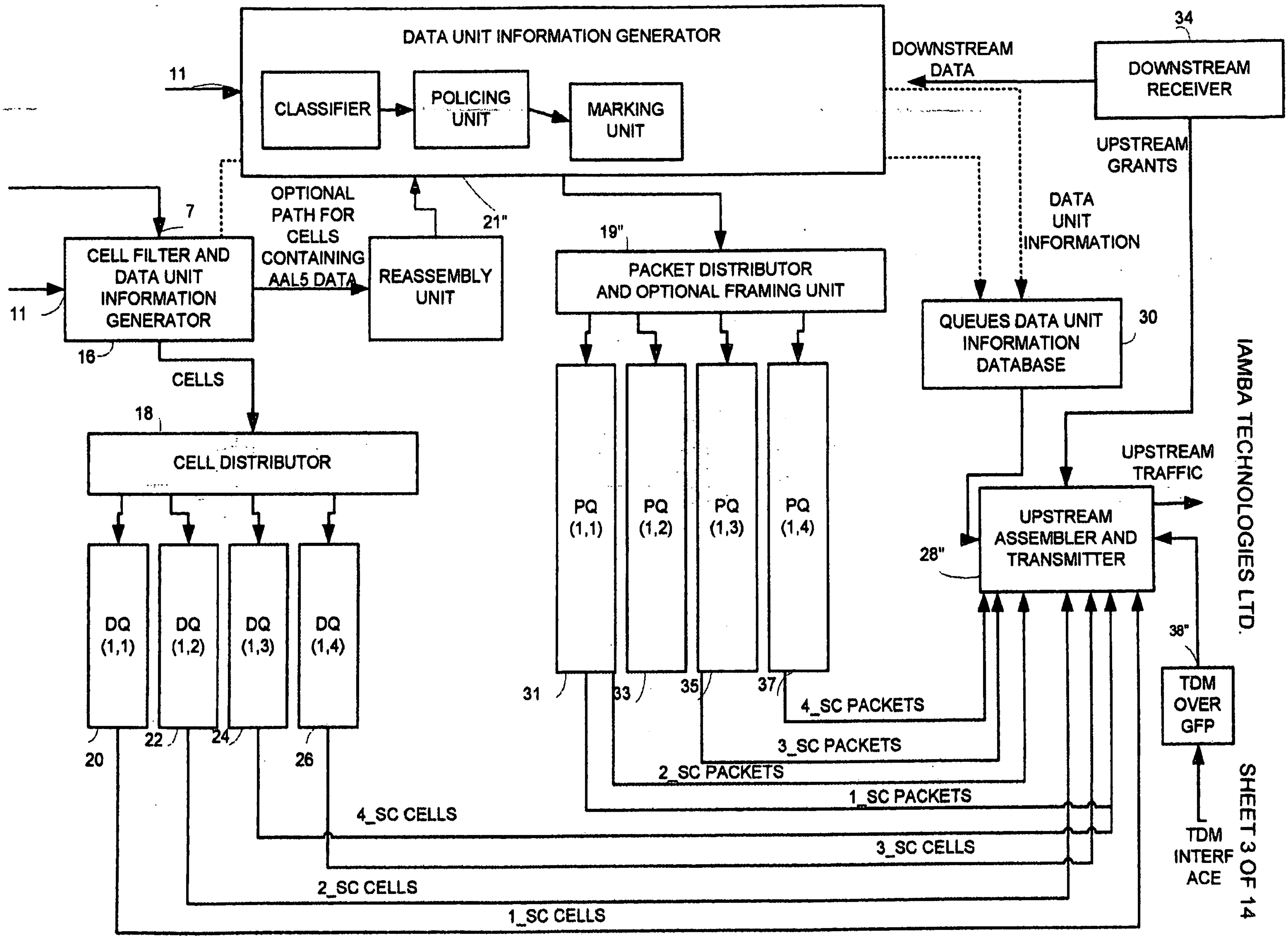
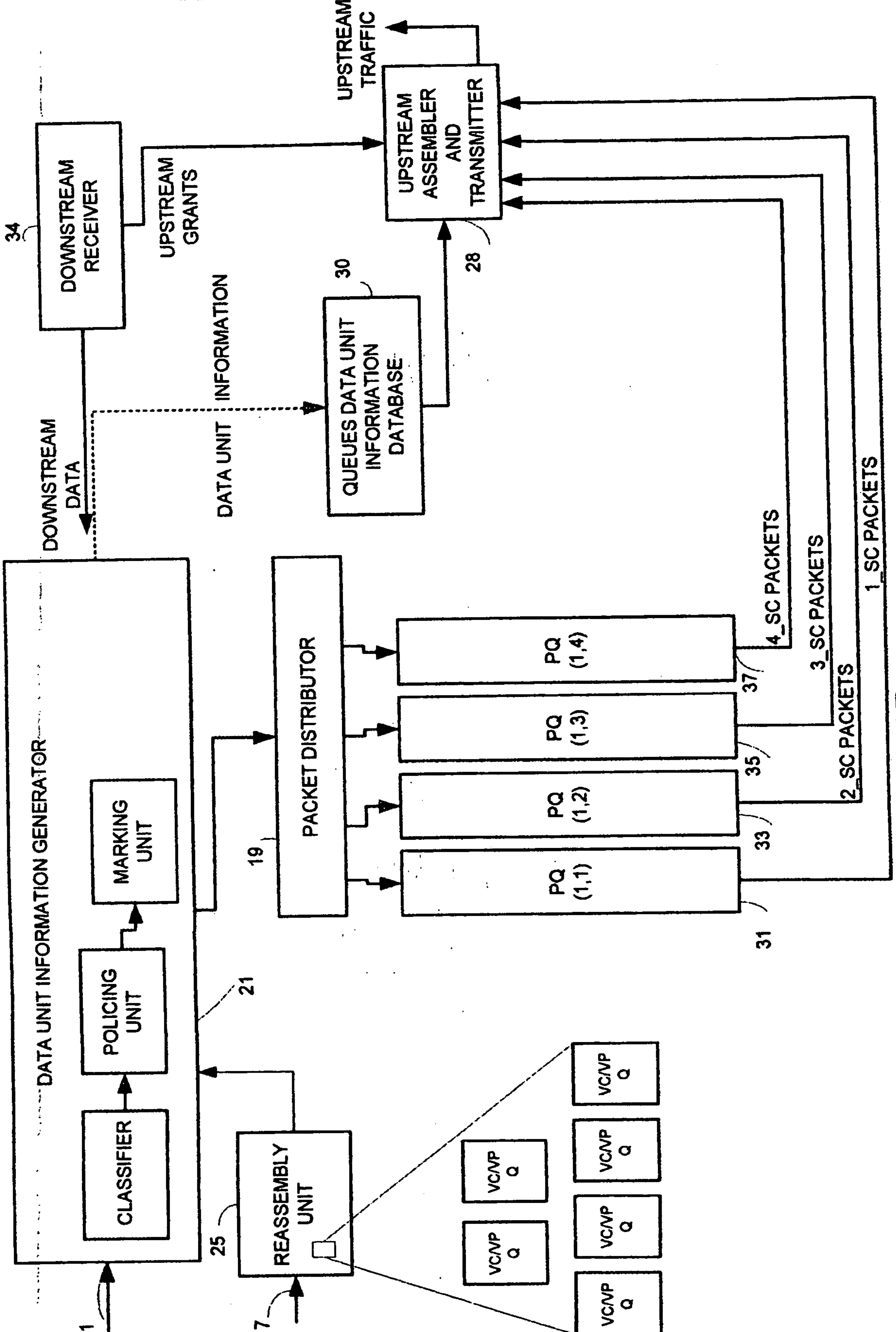

DATA UNIT INFORMATION GENERATOR
CLASSIFIER
POLICING UNIT
MARKING UNIT
11
7
21"
OPTIONAL PATH FOR CELLS CONTAINING AAL5 DATA
CELL FILTER AND DATA UNIT INFORMATION GENERATOR
REASSEMBLY UNIT
16
CELLS
18
CELL DISTRIBUTOR
DQ (1,1)
DQ (1,2)
DQ (1,3)
DQ (1,4)
20
22
24
26
19"
PACKET DISTRIBUTOR AND OPTIONAL FRAMING UNIT
PQ (1,1)
PQ (1,2)
PQ (1,3)
PQ (1,4)
31
33
35
37
4_SC PACKETS
3_SC PACKETS
2_SC PACKETS
1_SC PACKETS
4_SC CELLS
3_SC CELLS
2_SC CELLS
1_SC CELLS
DOWNSTREAM DATA
34
DOWNSTREAM RECEIVER
UPSTREAM GRANTS
DATA UNIT INFORMATION
QUEUES DATA UNIT INFORMATION DATABASE
30
UPSTREAM TRAFFIC
UPSTREAM ASSEMBLER AND TRANSMITTER
28"
38"
TDM OVER GFP
TDM INTERF ACE

FIG. 1C 9"



UPSTREAM TRAFFIC
DOWNSTREAM RECEIVER
34
UPSTREAM GRANTS
UPSTREAM ASSEMBLER AND TRANSMITTER
28
QUEUES DATA UNIT INFORMATION DATABASE
30
DATA UNIT INFORMATION
DOWNSTREAM DATA
DATA UNIT INFORMATION GENERATOR
CLASSIFIER
POLICING UNIT
MARKING UNIT
21
11
PACKET DISTRIBUTOR
19
PQ (1,1)
PQ (1,2)
PQ (1,3)
PQ (1,4)
31
33
35
37
1_SC PACKETS
2_SC PACKETS
3_SC PACKETS
4_SC PACKETS
REASSEMBLY UNIT
25
7
VC/VP Q
VC/VP Q
VC/VP Q
VC/VP Q
VC/VP Q
VC/VP Q

FIG. 1D

9*

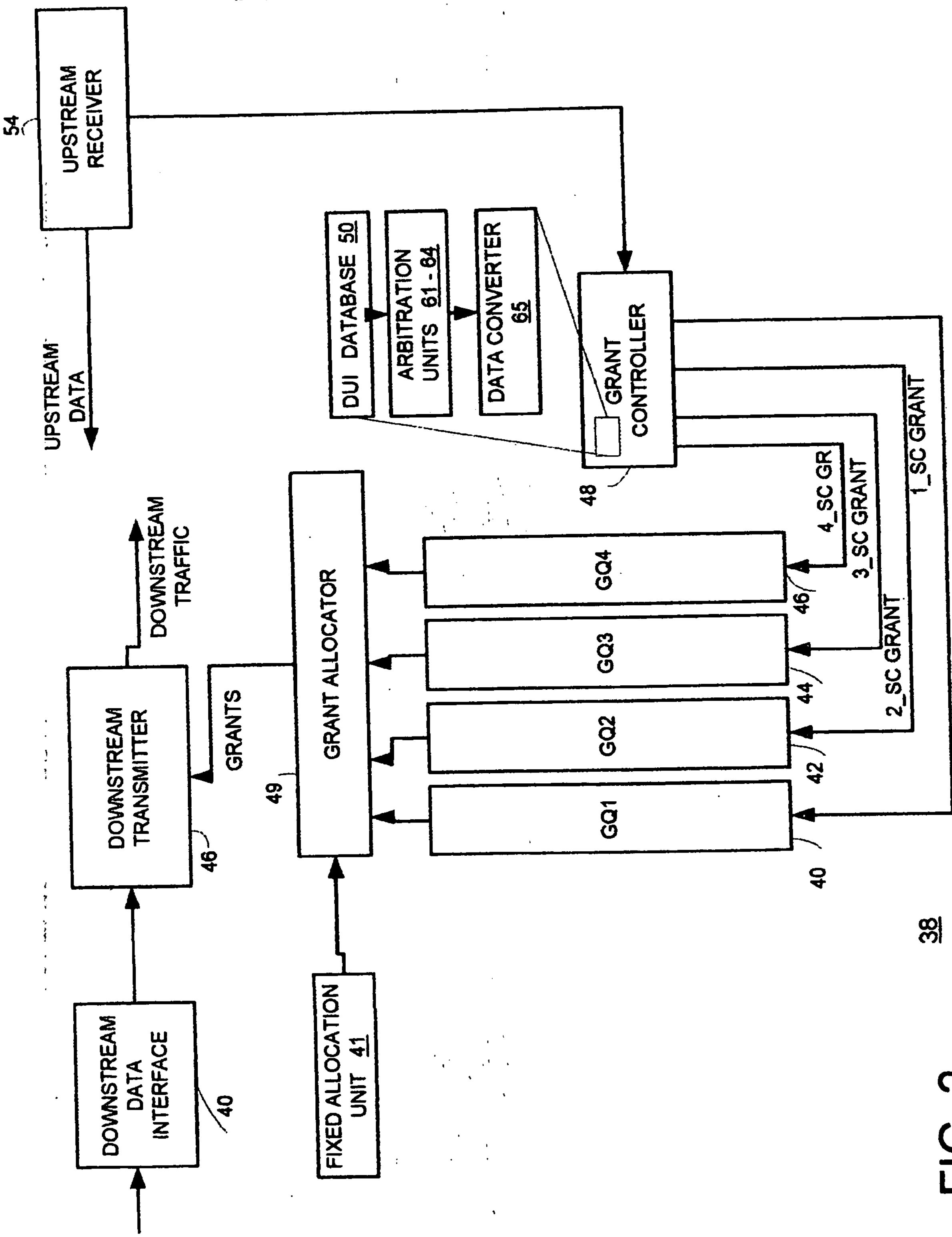

UPSTREAM RECEIVER
54
UPSTREAM DATA
DUI DATABASE 50
ARBITRATION UNITS 61 - 64
DATA CONVERTER 65
GRANT CONTROLLER
48
4_SC GR
3_SC GRANT
2_SC GRANT
1_SC GRANT
DOWNSTREAM TRAFFIC
GRANT ALLOCATOR
49
GQ4
GQ3
GQ2
GQ1
46
44
42
40
DOWNSTREAM TRANSMITTER
46
GRANTS
DOWNSTREAM DATA INTERFACE
40
FIXED ALLOCATION UNIT 41
38

FIG. 2

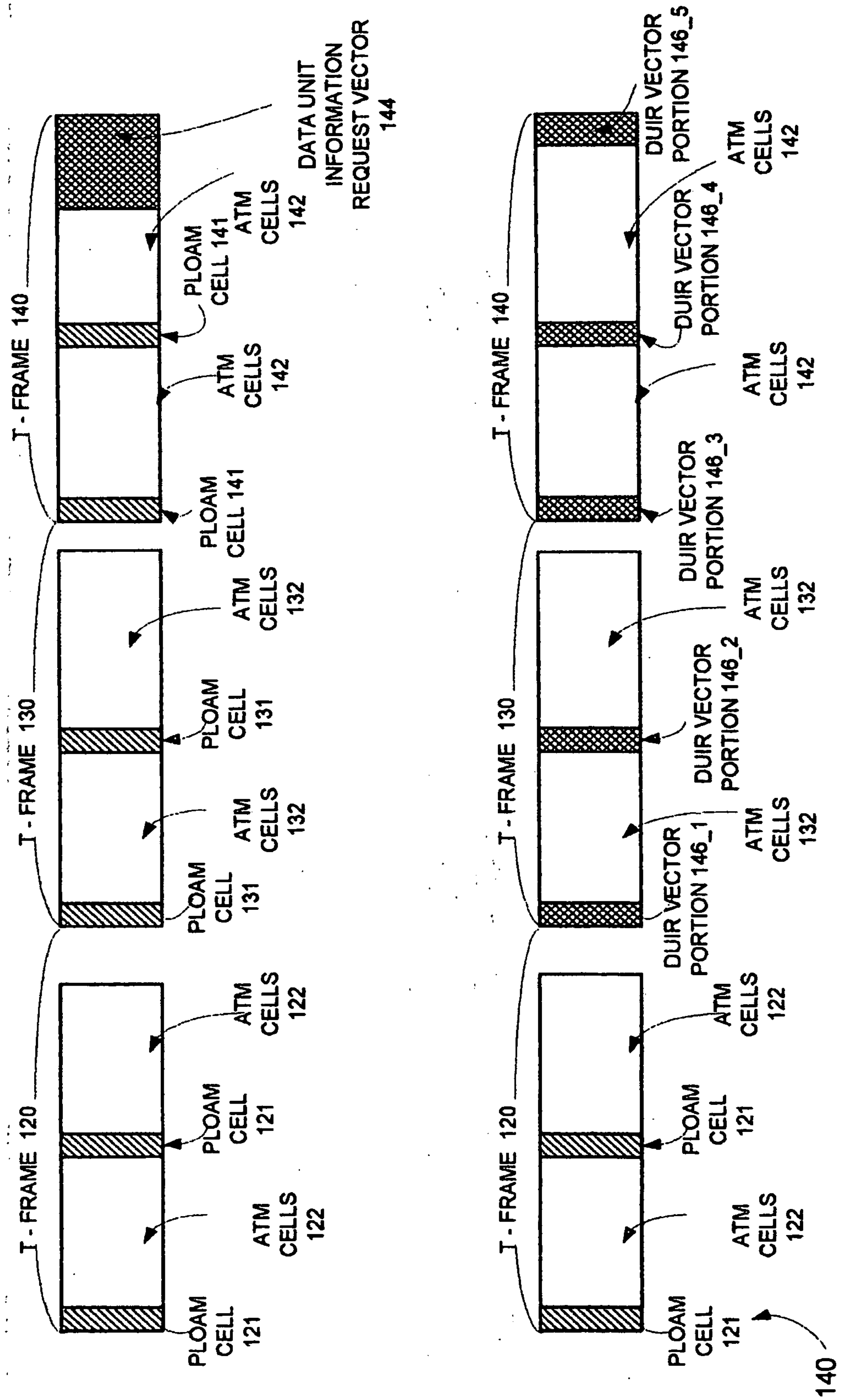
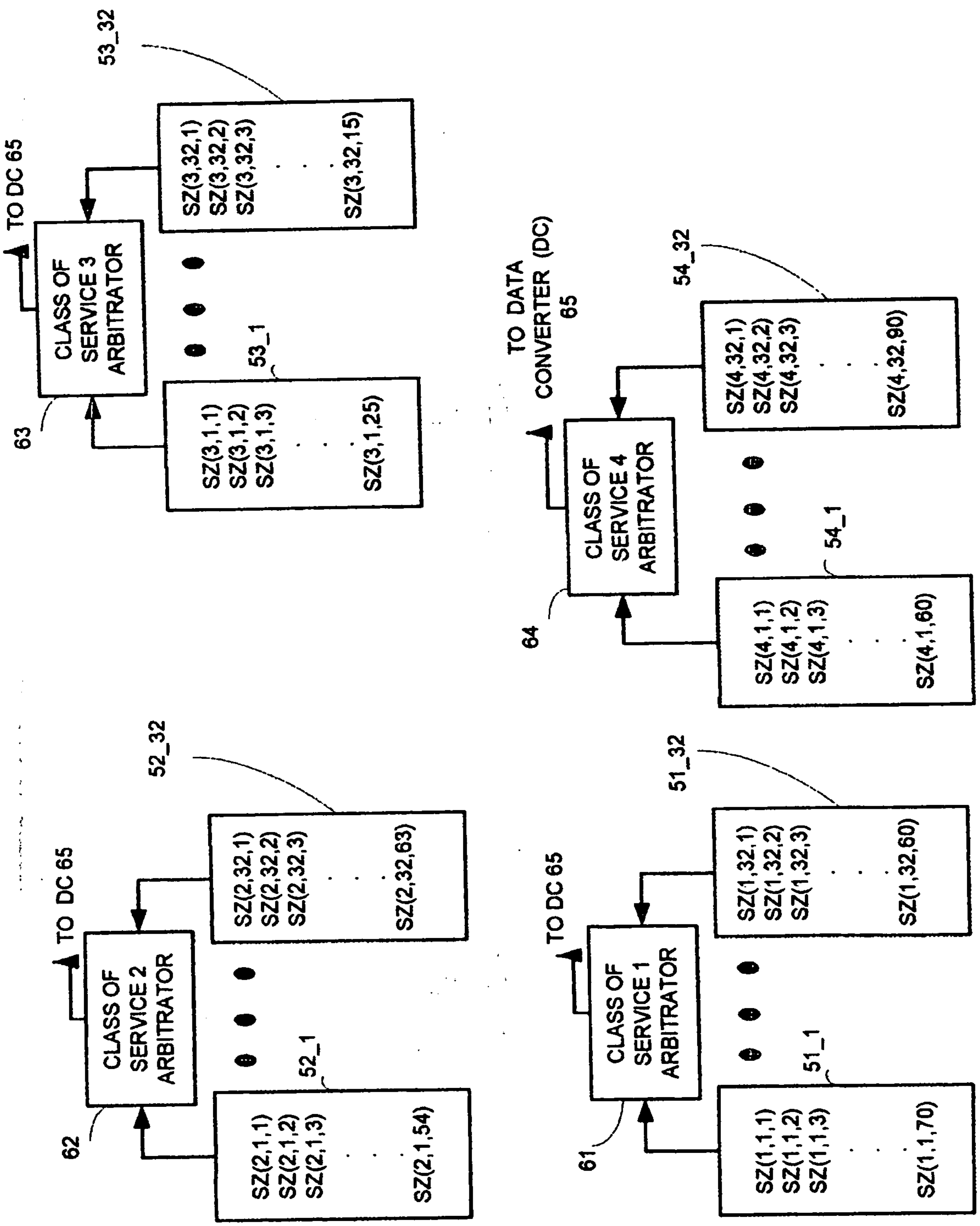

T - FRAME 120
T - FRAME 130
T - FRAME 140
PLOAM CELL 121
ATM CELLS 122
PLOAM CELL 131
ATM CELLS 132
PLOAM CELL 141
ATM CELLS 142
DATA UNIT INFORMATION REQUEST VECTOR 144
DUIR VECTOR PORTION 146_1
DUIR VECTOR PORTION 146_2
DUIR VECTOR PORTION 146_3
DUIR VECTOR PORTION 146_4
DUIR VECTOR PORTION 146_5
140

FIG. 3B



CLASS OF SERVICE 1 ARBITRATOR
61
TO DC 65
SZ(1,1,1)
SZ(1,1,2)
SZ(1,1,3)
SZ(1,1,70)
51_1
SZ(1,32,1)
SZ(1,32,2)
SZ(1,32,3)
SZ(1,32,60)
51_32
CLASS OF SERVICE 2 ARBITRATOR
62
TO DC 65
SZ(2,1,1)
SZ(2,1,2)
SZ(2,1,3)
SZ(2,1,54)
52_1
SZ(2,32,1)
SZ(2,32,2)
SZ(2,32,3)
SZ(2,32,63)
52_32
CLASS OF SERVICE 3 ARBITRATOR
63
TO DC 65
SZ(3,1,1)
SZ(3,1,2)
SZ(3,1,3)
SZ(3,1,25)
53_1
SZ(3,32,1)
SZ(3,32,2)
SZ(3,32,3)
SZ(3,32,15)
53_32
CLASS OF SERVICE 4 ARBITRATOR
64
TO DATA CONVERTER (DC) 65
SZ(4,1,1)
SZ(4,1,2)
SZ(4,1,3)
SZ(4,1,60)
54_1
SZ(4,32,1)
SZ(4,32,2)
SZ(4,32,3)
SZ(4,32,90)
54_32

FIG. 4

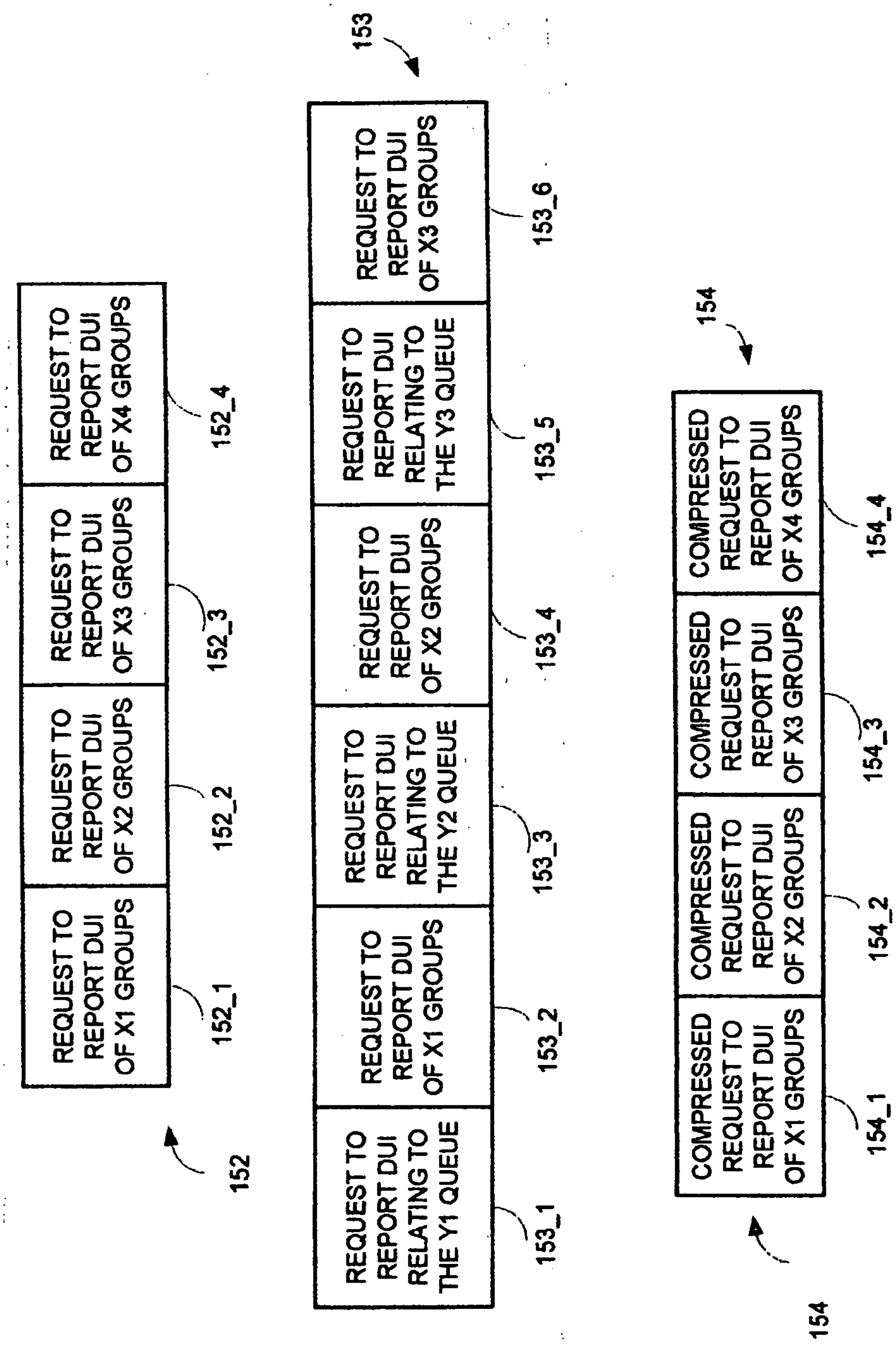

152
REQUEST TO REPORT DUI OF X1 GROUPS
152_1
REQUEST TO REPORT DUI OF X2 GROUPS
152_2
REQUEST TO REPORT DUI OF X3 GROUPS
152_3
REQUEST TO REPORT DUI OF X4 GROUPS
152_4
153
REQUEST TO REPORT DUI RELATING TO THE Y1 QUEUE
153_1
REQUEST TO REPORT DUI OF X1 GROUPS
153_2
REQUEST TO REPORT DUI RELATING TO THE Y2 QUEUE
153_3
REQUEST TO REPORT DUI OF X2 GROUPS
153_4
REQUEST TO REPORT DUI RELATING TO THE Y3 QUEUE
153_5
REQUEST TO REPORT DUI OF X3 GROUPS
153_6
154
COMPRESSED REQUEST TO REPORT DUI OF X1 GROUPS
154_1
COMPRESSED REQUEST TO REPORT DUI OF X2 GROUPS
154_2
COMPRESSED REQUEST TO REPORT DUI OF X3 GROUPS
154_3
COMPRESSED REQUEST TO REPORT DUI OF X4 GROUPS
154_4

FIG. 5A

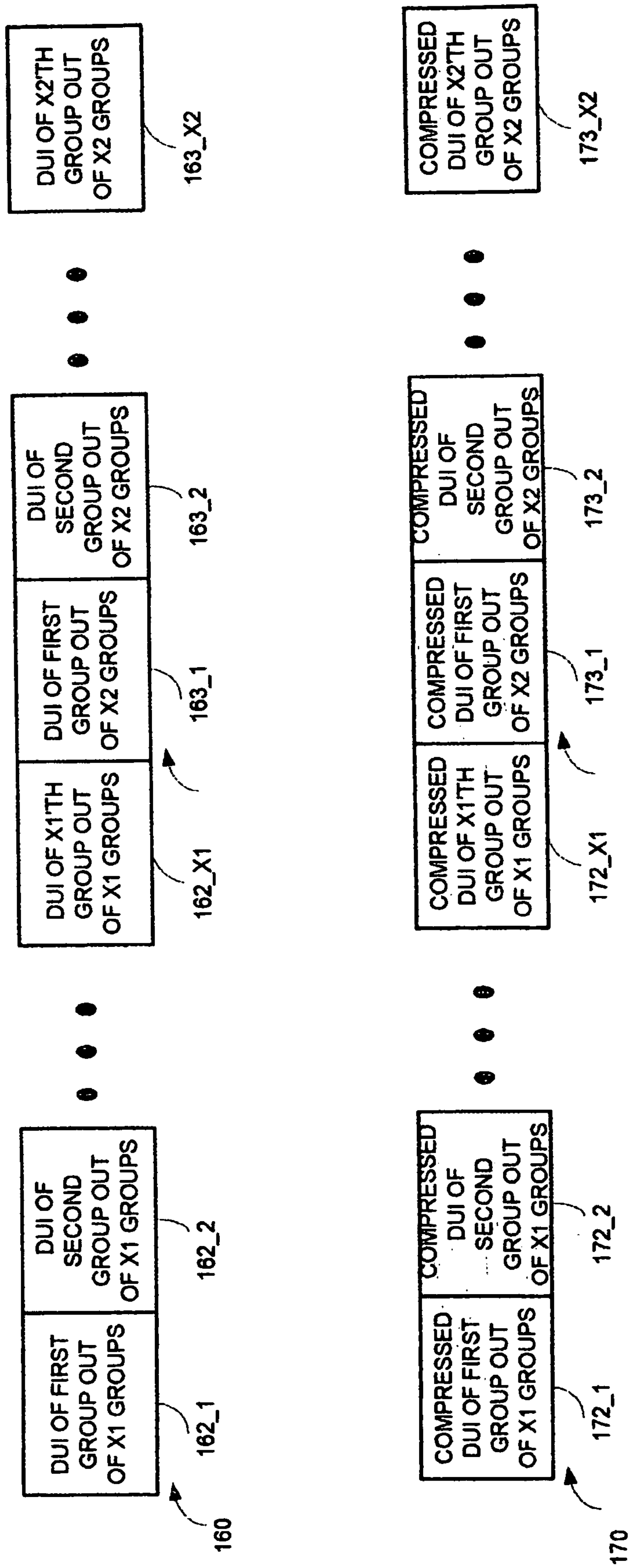

160
DUI OF FIRST GROUP OUT OF X1 GROUPS
162_1
DUI OF SECOND GROUP OUT OF X1 GROUPS
162_2
DUI OF X1'TH GROUP OUT OF X1 GROUPS
162_X1
DUI OF FIRST GROUP OUT OF X2 GROUPS
163_1
DUI OF SECOND GROUP OUT OF X2 GROUPS
163_2
DUI OF X2'TH GROUP OUT OF X2 GROUPS
163_X2
170
COMPRESSED DUI OF FIRST GROUP OUT OF X1 GROUPS
172_1
COMPRESSED DUI OF SECOND GROUP OUT OF X1 GROUPS
172_2
COMPRESSED DUI OF X1'TH GROUP OUT OF X1 GROUPS
172_X1
COMPRESSED DUI OF FIRST GROUP OUT OF X2 GROUPS
173_1
COMPRESSED DUI OF SECOND GROUP OUT OF X2 GROUPS
173_2
COMPRESSED DUI OF X2'TH GROUP OUT OF X2 GROUPS
173_X2

FIG. 5B

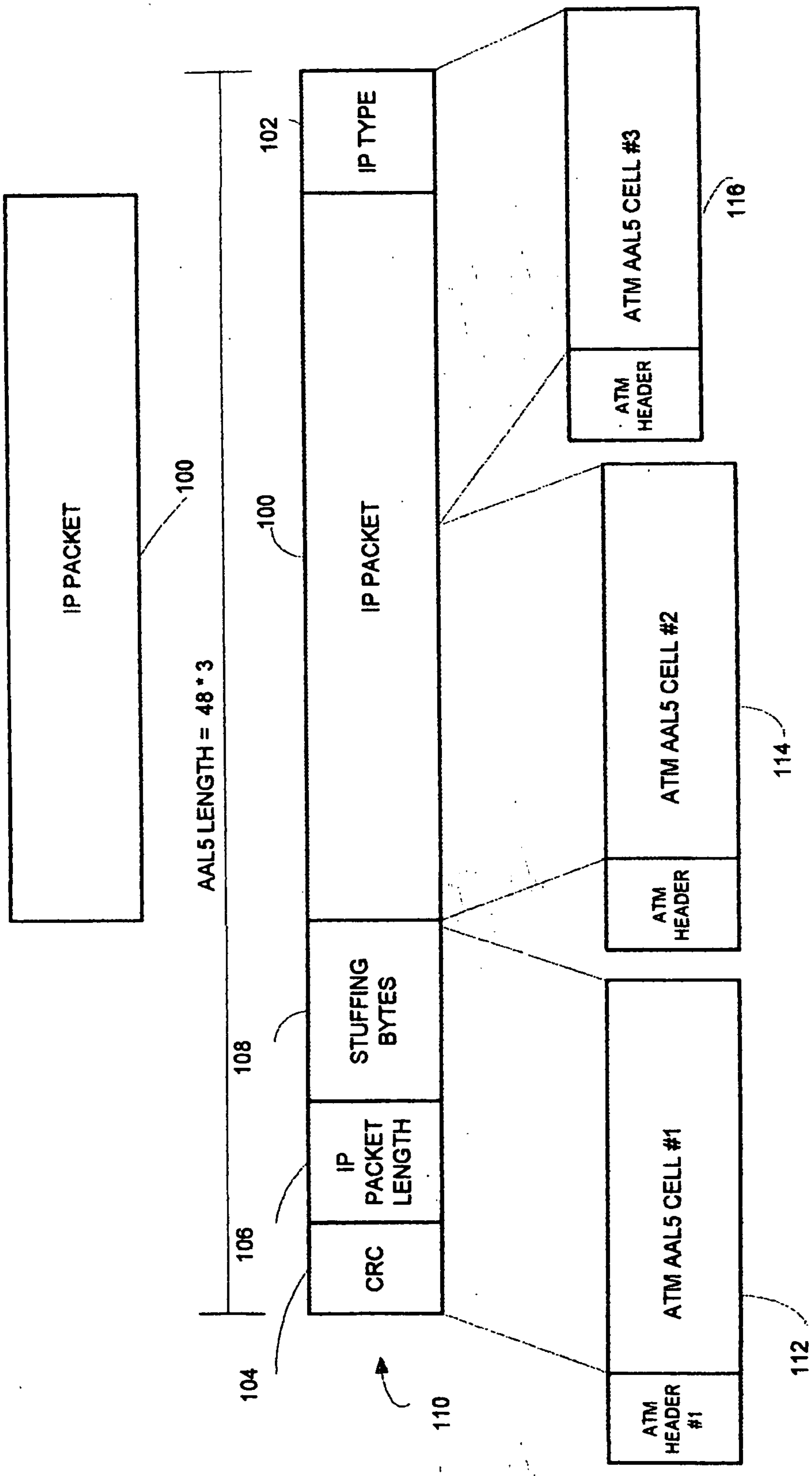

IP PACKET
100
AAL5 LENGTH = 48 * 3
102
IP TYPE
100
IP PACKET
108
STUFFING BYTES
106
IP PACKET LENGTH
104
CRC
110
ATM AAL5 CELL #3
ATM HEADER
116
ATM AAL5 CELL #2
ATM HEADER
114
ATM AAL5 CELL #1
ATM HEADER #1
112

FIG. 3A

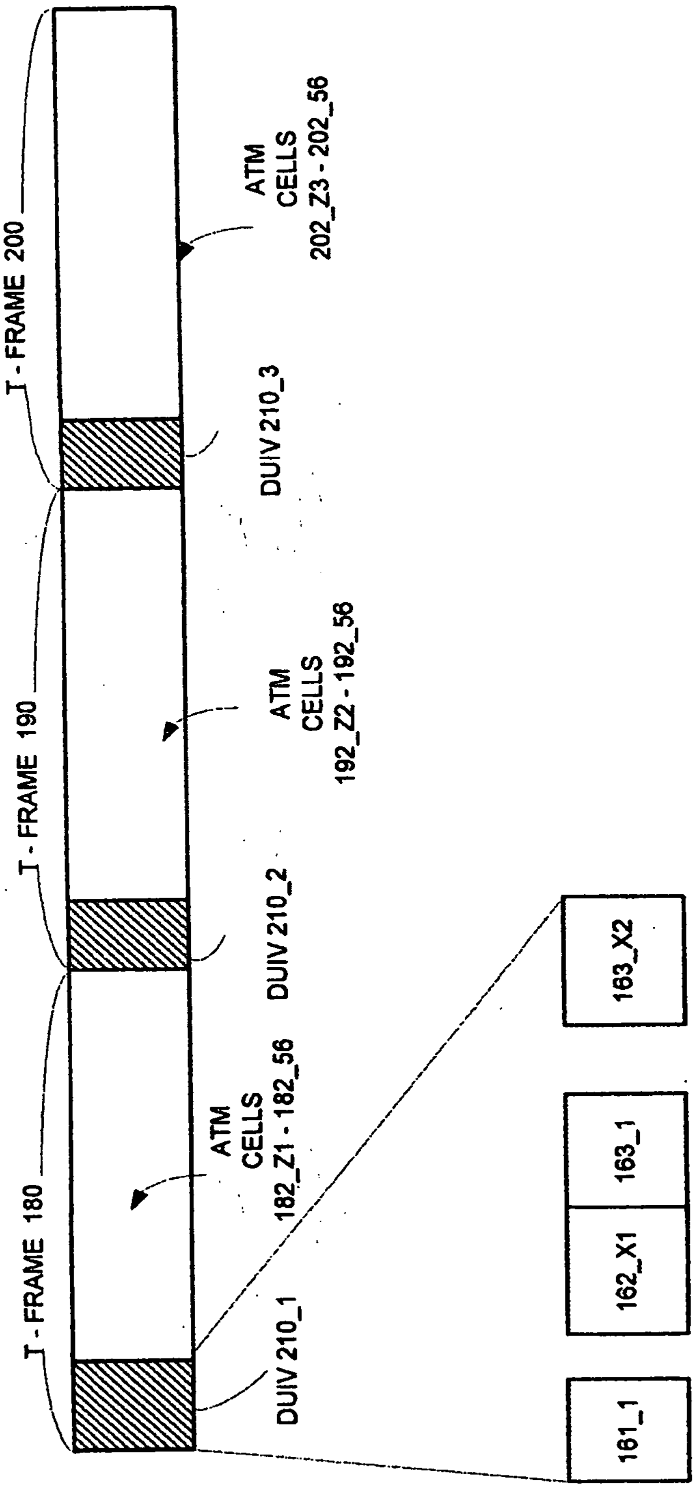

T - FRAME 200
T - FRAME 190
T - FRAME 180
ATM CELLS 202_Z3 - 202_56
ATM CELLS 192_Z2 - 192_56
ATM CELLS 182_Z1 - 182_56
DUIV 210_3
DUIV 210_2
DUIV 210_1
163_X2
163_1
162_X1
161_1

FIG. 5C

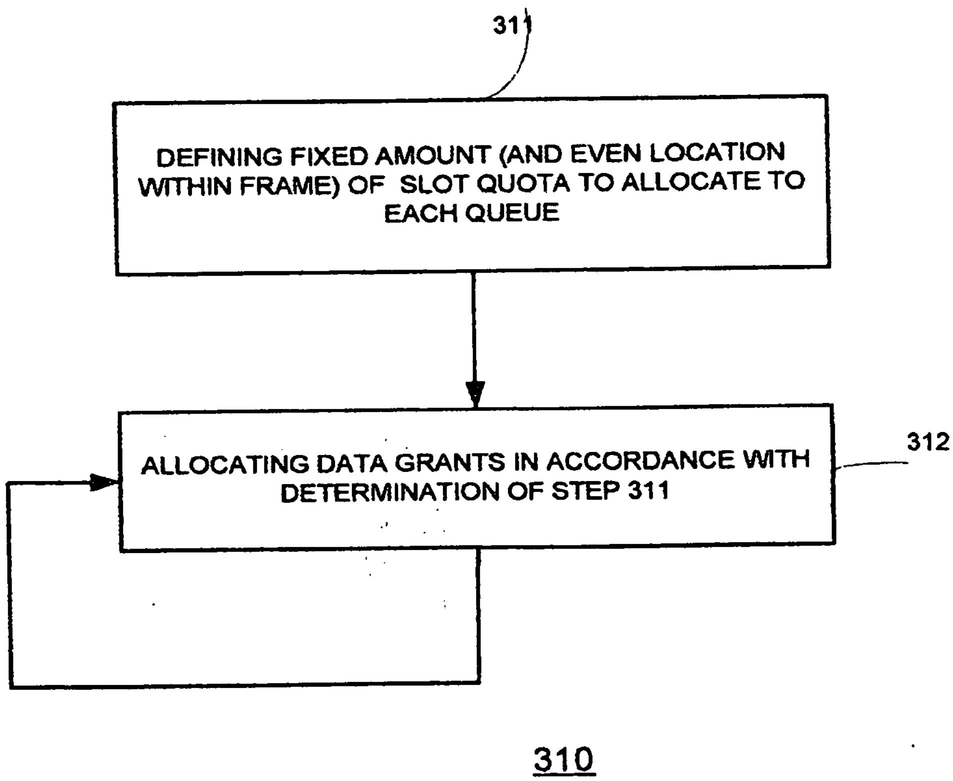
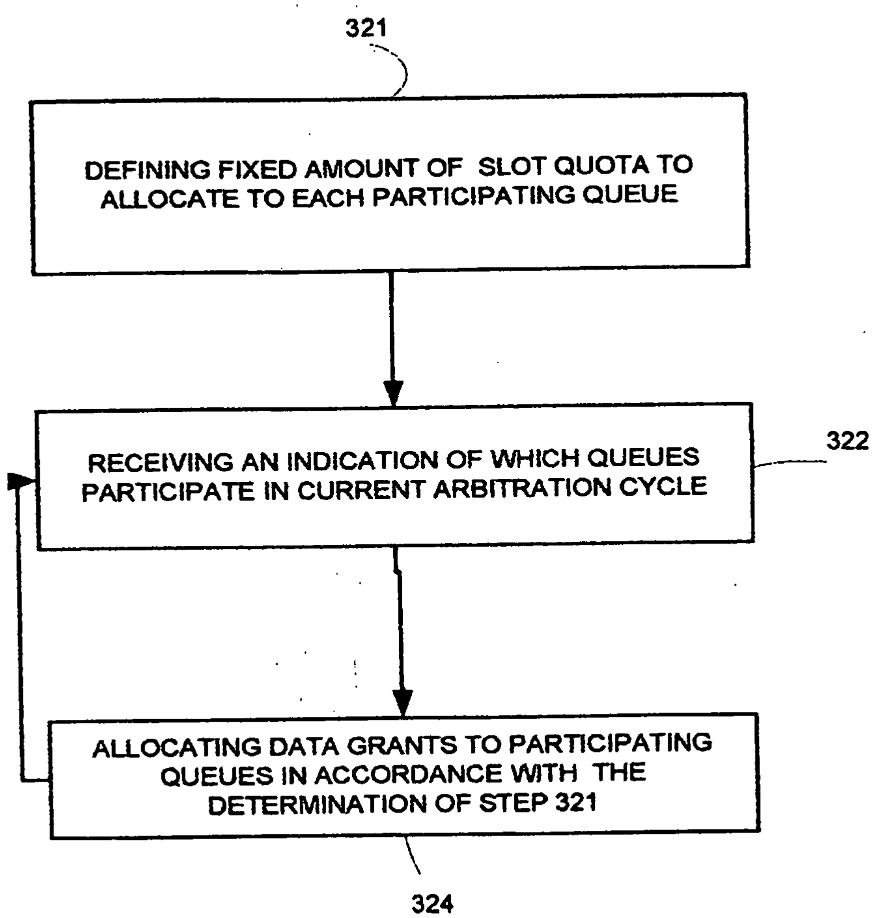

311
DEFINING FIXED AMOUNT (AND EVEN LOCATION WITHIN FRAME) OF SLOT QUOTA TO ALLOCATE TO EACH QUEUE
ALLOCATING DATA GRANTS IN ACCORDANCE WITH DETERMINATION OF STEP 311
312
310

FIG. 6A


321
DEFINING FIXED AMOUNT OF SLOT QUOTA TO ALLOCATE TO EACH PARTICIPATING QUEUE
RECEIVING AN INDICATION OF WHICH QUEUES PARTICIPATE IN CURRENT ARBITRATION CYCLE
322
ALLOCATING DATA GRANTS TO PARTICIPATING QUEUES IN ACCORDANCE WITH THE DETERMINATION OF STEP 321
324

320

FIG. 6B

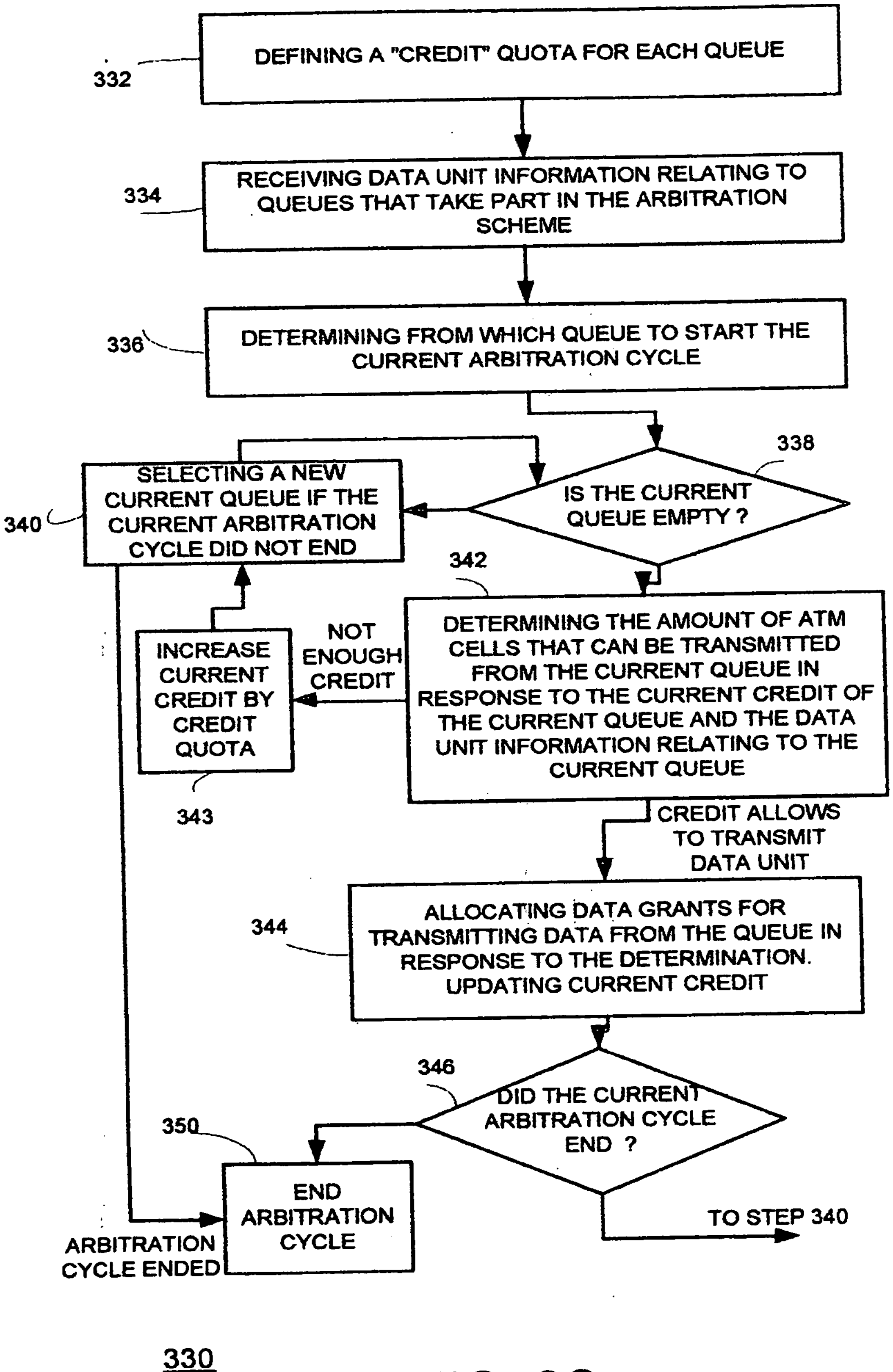

332
DEFINING A "CREDIT" QUOTA FOR EACH QUEUE
334
RECEIVING DATA UNIT INFORMATION RELATING TO QUEUES THAT TAKE PART IN THE ARBITRATION SCHEME
336
DETERMINING FROM WHICH QUEUE TO START THE CURRENT ARBITRATION CYCLE
338
IS THE CURRENT QUEUE EMPTY ?
340
SELECTING A NEW CURRENT QUEUE IF THE CURRENT ARBITRATION CYCLE DID NOT END
342
DETERMINING THE AMOUNT OF ATM CELLS THAT CAN BE TRANSMITTED FROM THE CURRENT QUEUE IN RESPONSE TO THE CURRENT CREDIT OF THE CURRENT QUEUE AND THE DATA UNIT INFORMATION RELATING TO THE CURRENT QUEUE
NOT ENOUGH CREDIT
INCREASE CURRENT CREDIT BY CREDIT QUOTA
343
CREDIT ALLOWS TO TRANSMIT DATA UNIT
344
ALLOCATING DATA GRANTS FOR TRANSMITTING DATA FROM THE QUEUE IN RESPONSE TO THE DETERMINATION. UPDATING CURRENT CREDIT
346
DID THE CURRENT ARBITRATION CYCLE END ?
350
END ARBITRATION CYCLE
ARBITRATION CYCLE ENDED
TO STEP 340
330

FIG. 6C